

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 24, 2003, 18:40:04 ; Search time 2232 Seconds  
(without alignments)  
123.353 Million cell updates/sec

Title: US-09-771-009-1  
Perfect score: 90  
Sequence: 1 NFDLLKLAGDVESNPGP 17

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n\_model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spo01/US09771009/runat\_22012003\_140101\_18469/app.query.fasta\_1.199  
-DB=EST -QFMT=fastp -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plo -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09771009@cgn.1.1.1716 -runat\_22012003\_140101\_18469 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_estl.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
C 1	57	63.3	758	17	A0653827	A0653827	Sheared D
C 2	56	62.2	536	17	A0651117	A0651117	Sheared D
C 3	54	60.0	369	17	AQ03690	GSSTC0473	AQ03690 GSSTC0473
C 4	54	60.0	511	17	A2157335	SP_0024_A	A2157335 SP_0024_A
C 5	54	60.0	516	17	A2157346	SP_0024_A	A2157346 SP_0024_A
C 6	54	60.0	522	17	A2156841	SP_0049_B	A2156841 SP_0049_B
C 7	54	60.0	573	17	A0653839	927P1-15E	A0653839 927P1-15E
C 8	53.5	59.4	781	17	A2194671	SP_1028_A	A2194671 SP_1028_A
C 9	53	58.9	288	17	AZ302518	GSSTC1239	AZ302518 GSSTC1239
C 10	53	58.9	326	17	AQ094204	GSSTC0594	AQ094204 GSSTC0594
C 11	53	58.9	326	17	AQ0908490	GSSTC0544	AQ0908490 GSSTC0544
C 12	53	58.9	359	17	AZ302155	GSSTC1226	AZ302155 GSSTC1226
C 13	53	58.9	379	17	AQ445296	GSSTC0157	AQ445296 GSSTC0157
C 14	53	58.9	418	17	AQ0910615	GSSTC0345	AQ0910615 GSSTC0345
C 15	53	58.9	427	17	AZ302187	GSSTC1223	AZ302187 GSSTC1223
C 16	53	58.9	451	17	BH842561	TC3-53J14	BH842561 TC3-53J14
C 17	53	58.9	472	17	BH192203	TC3-68G19	BH192203 TC3-68G19
C 18	53	58.9	501	17	BH844419	TC3-58K4	BH844419 TC3-58K4
C 19	53	58.9	518	17	BH194877	TC3-71F5	BH194877 TC3-71F5
C 20	53	58.9	597	17	TA195C08P	SP_1004_B	TA195C08P SP_1004_B
C 21	53	58.9	817	17	AZ185217	T. brucei	AZ185217 T. brucei
C 22	52.5	58.3	526	17	TA361D08P	T. brucei	TA361D08P T. brucei
C 23	52	57.8	177	17	AZ218377	Sheared D	AZ218377 Sheared D
C 24	52	57.8	244	17	AQ940096	Sheared D	AQ940096 Sheared D
C 25	52	57.8	250	17	AQ659566	Sheared D	AQ659566 Sheared D
C 26	52	57.8	262	17	AQ655191	Sheared D	AQ655191 Sheared D
C 27	52	57.8	267	17	TA235C06Q	T. brucei	TA235C06Q T. brucei
C 28	52	57.8	291	17	AQ950047	Sheared D	AQ950047 Sheared D
C 29	52	57.8	302	17	TA312F02Q	T. brucei	TA312F02Q T. brucei
C 30	52	57.8	303	17	TA81E01Q	T. brucei	TA81E01Q T. brucei
C 31	52	57.8	312	17	AQ949973	Sheared D	AQ949973 Sheared D
C 32	52	57.8	335	17	AQ639202	927P1-17E	AQ639202 927P1-17E
C 33	52	57.8	336	17	AZ212305	Sheared D	AZ212305 Sheared D
C 34	52	57.8	342	17	TA64H01Q	T. brucei	TA64H01Q T. brucei
C 35	52	57.8	361	17	AZ218893	Sheared D	AZ218893 Sheared D
C 36	52	57.8	383	17	AZ218882	Sheared D	AZ218882 Sheared D
C 37	52	57.8	385	17	TA375A02P	T. brucei	TA375A02P T. brucei
C 38	52	57.8	400	17	AZ219198	Sheared D	AZ219198 Sheared D
C 39	52	57.8	401	17	AQ640112	927P1-2D1	AQ640112 927P1-2D1
C 40	52	57.8	403	17	AQ940610	Sheared D	AQ940610 Sheared D
C 41	52	57.8	407	17	AQ953669	Sheared D	AQ953669 Sheared D
C 42	52	57.8	413	17	AQ657002	Sheared D	AQ657002 Sheared D
C 43	52	57.8	415	17	AQ941204	Sheared D	AQ941204 Sheared D
C 44	52	57.8	424	17	AQ647225	RPC193-Dp	AQ647225 RPC193-Dp
C 45	52	57.8	427	17	AQ641585	RPC193-Ec	AQ641585 RPC193-Ec

ALIGNMENTS

RESULT 1  
AQ653827/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AQ653827  
Sheared DNA-1E19-TF Sheared DNA  
758 bp  
linear  
GSS 22-JUN-1999  
Sheared DNA-1E19, DNA sequence.  
AQ653827.1  
GI:5147013  
GSS.  
Trypanosoma brucei.  
Trypanosoma brucei.  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma.  
1 (bases 1 to 758)  
El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Suh, E., Malek, J., Fujii, C.,  
Gerrard, C., Leech, V., de Jong, P., Ullu, E., Melville, S., Donelson, J.



Tel: 54-11-4580-7255 ext 309  
 Fax: 54-11-4752-9639  
 Email: dsanchez@lib.unsam.edu.ar  
 Sequences were basecalled with phred and vector was masked with  
 crossmatch (see http://genome.washington.edu). Sequences were then  
 trimmed from both ends to remove low quality bases and masked  
 vector.  
 Seq primer: T7  
 Class: shotgun.

## FEATURES

source  
 Location/Qualifiers  
 1..369  
 /organism="Trypanosoma cruzi"  
 /strain="CL-Brener"  
 /db\_xref="taxon:5693"  
 /clone="G23122"  
 /clone\_lib="Trypanosoma cruzi random genomic library"  
 /cell\_type="Epimastigote"  
 /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was  
 randomly sheared using a nebulizer and the 1 to 2 Kb range  
 was gel purified and cloned into the dephosphorylated  
 HincII site of the vector."  
 BASE COUNT 93 a 101 c 100 g 75 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 34.3 Length: 369  
 Score: 54.00 Matches: 10  
 Percent Similarity: 84.62% Conservative: 1  
 Best Local Similarity: 76.92% Mismatches: 2  
 Query Match: 60.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-771-009-1 (1-17) x A0903690 (1-369)

QY 5 LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17

Db 62 CTACTGCTAGCGAGACGCTTGAGCAGACCCAGGCCCC 100

RESULT 4  
 AZ157335  
 LOCUS  
 DEFINITION  
 SP\_0024\_AL\_B05\_T7 Strongylocentrotus purpuratus, purple sea urchin,  
 sperm genomic BAC library Strongylocentrotus purpuratus genomic  
 clone Plate=24 Col=9 Row=C, DNA sequence.

## ACCESSION

VERSION  
 AZ157335

## KEYWORDS

SOURCE  
 ORGANISM

Strongylocentrotus purpuratus.  
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoida;  
 Strongylocentrotidae; Strongylocentrotus.

## REFERENCE

AUTHORS  
 Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,  
 Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray  
 ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and  
 Hood,L.

TITLE  
 A sea urchin genome project: Sequence scan, virtual map, and  
 additional resources

JOURNAL  
 MEDLINE  
 COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

Contact: Cameron, RA, Davidson, EH, Hood, L

Division of Biology 156-29

California Institute of Technology

Pasadena California 91125, USA

Tel: (626) 395-8421

Fax: (626) 793-3047

Email: acameron@caltech.edu

Plate: 24 row: C column: 9

Seq primer: T7

Class: BAC ends

High quality sequence stop: 511.

Location/Qualifiers

## FEATURES

Alignment Scores:

Pred. No.: 50.2 Length: 516

## source

1..511  
 /organism="Strongylocentrotus purpuratus"  
 /db\_xref="taxon:7668"  
 /clone="Plate=24 Col=9 Row=C"  
 /clone\_lib="Strongylocentrotus purpuratus, purple sea  
 urchin, sperm genomic BAC library"  
 /note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli  
 DH10B"

BASE COUNT 140 a 104 c 101 g 164 t 2 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 49.6 Length: 511  
 Score: 54.00 Matches: 10  
 Percent Similarity: 85.71% Conservative: 2  
 Best Local Similarity: 71.43% Mismatches: 2  
 Query Match: 60.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-771-009-1 (1-17) x AZ157335 (1-511)

QY 4 LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17

Db 323 CTCCTGATGATCTGCTGACGTGAACCAATCCAGGTCCT 364

## RESULT 5

AZ157346

LOCUS

DEFINITION

SP\_0024\_AL\_C05\_T7 Strongylocentrotus purpuratus, purple sea urchin,  
 sperm genomic BAC library Strongylocentrotus purpuratus genomic  
 clone Plate=24 Col=9 Row=E, DNA sequence.

ACCESSION

VERSION

AZ157346

KEYWORDS

SOURCE

ORGANISM

Strongylocentrotus purpuratus.

Strongylocentrotus purpuratus

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

Echinoidea; Euechinoidea; Echinacea; Echinoida;

Strongylocentrotidae; Strongylocentrotus.

1 (bases 1 to 516)

AUTHORS

Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,  
 Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray  
 ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and  
 Hood,L.

TITLE

A sea urchin genome project: Sequence scan, virtual map, and  
 additional resources

JOURNAL

MEDLINE

COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

Contact: Cameron, RA, Davidson, EH, Hood, L

Division of Biology 156-29

California Institute of Technology

Pasadena California 91125, USA

Tel: (626) 395-8421

Fax: (626) 793-3047

Email: acameron@caltech.edu

Plate: 24 row: E column: 9

Seq primer: T7

Class: BAC ends

High quality sequence stop: 516.

Location/Qualifiers

1..516

/organism="Strongylocentrotus purpuratus"

/db\_xref="taxon:7668"

/clone="Plate=24 Col=9 Row=E"

/clone\_lib="Strongylocentrotus purpuratus, purple sea  
 urchin, sperm genomic BAC library"

/note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli  
 DH10B"

BASE COUNT 143 a 104 c 98 g 168 t 3 others  
 ORIGIN

Alignment Scores:

Pred. No.: 50.2 Length: 516

Score: 54.00 Matches: 10  
 Percent Similarity: 85.71% Conservative: 2  
 Best Local Similarity: 71.43% Mismatches: 2  
 Query Match: 60.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-771-009-1 (1-17) x AZ157346 (1-516)

Qy 4 LeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 A2156841  
 DB 326 CTCCTGATGAGATCTGGTACGTTGAACCAATCCAGGTCCT 367  
 ||||| :|||:|||||:|||||:|||||:|||||

RESULT 6  
 A2156841  
 LOCUS SP\_0049\_B2\_G05\_T7 522 bp DNA linear GSS 29-AUG-2000  
 DEFINITION Strongylocentrotus purpuratus, purple sea urchin,  
 sperm genomic BAC library Strongylocentrotus purpuratus genomic  
 clone Plate-49 Col=10 Row=N, DNA sequence.

ACCESSION A2156841  
 VERSION A2156841  
 KEYWORDS GSS.  
 SOURCE Strongylocentrotus purpuratus.  
 ORGANISM Strongylocentrotus purpuratus.

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoidea;  
 Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 522)  
 AUTHORS Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,  
 Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,  
 G.A., Ettensohn,C.A., Leirach,H., Britten,R.J., Davidson,E.H. and  
 Hood,L.

TITLE A sea urchin genome project: Sequence scan, virtual map, and  
 additional resources

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)  
 MEDLINE 20402566  
 COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L  
 Division of Biology 156-29  
 California Institute of Technology  
 Pasadena California 91125, USA

Tel: (626) 395-8421  
 Fax: (626) 793-3047  
 Email: acameron@caltech.edu  
 Plate: 49 row: N column: 10  
 Seq primer: T7  
 Class: BAC ends  
 High quality sequence stop: 522.

FEATURES Location/Qualifiers

1..522  
 /organism="Strongylocentrotus purpuratus"  
 /db\_xref="taxon:7668"  
 /clone="Plate-49 Col=10 Row=N"  
 /clone\_lib="Strongylocentrotus purpuratus, purple sea  
 urchin, sperm genomic BAC library"  
 /note="Organ: sperm; Vector: BACE3.6; BAC Clones in E-Coli  
 DH10B"

BASE COUNT 131 a 101 c 107 g 182 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 50.8 Length: 522  
 Score: 54.00 Matches: 10  
 Percent Similarity: 85.71% Conservative: 2  
 Best Local Similarity: 71.43% Mismatches: 2  
 Query Match: 60.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-771-009-1 (1-17) x A2156841 (1-522)

Qy 4 LeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 A2156841  
 DB 177 CTCCTGATGAGATCTGGTACGTTGAACCAACCTGGACCA 218  
 ||||| :|||:|||||:|||||:|||||:|||||

RESULT 7

AQ639839  
 LOCUS 527P1-15E10.TP 573 bp DNA linear GSS 08-JUL-1999  
 DEFINITION Trypanosoma brucei genomic clone 927P1-15E10,  
 DNA sequence.

ACCESSION AQ639839  
 VERSION AQ639839.1 GI:5116549  
 KEYWORDS GSS.  
 SOURCE Trypanosoma brucei.  
 ORGANISM Trypanosoma brucei.

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma.  
 1 (bases 1 to 573)  
 El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,  
 Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J.,  
 Fraser,C. and Adams,M.

TITLE Determination of clone end sequences from Trypanosoma brucei TREU  
 927/4 P1 library  
 JOURNAL Unpublished (1999)  
 COMMENT Other\_GSSs: 927P1-15E10.TV  
 Contact: Najib M. El-Sayed  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: nelsayed@tigr.org

For clone/filter availability, please contact Sara Melville  
 (sm160@olebio.cam.ac.uk). P1 end sequences search page:  
 http://www.tigr.org/tdb/mdb/tbdb/.

Seq primer: SP6

Class: P1 ends.

FEATURES Location/Qualifiers

1..573  
 /organism="Trypanosoma brucei"  
 /strain="TREU927/4"  
 /db\_xref="taxon:5691"  
 /clone="927P1-15E10"  
 /clone\_lib="927P1"  
 /note="Vector: pAD10SacBII; Site1: Bam HI; Constructed by  
 Sara Melville, University of Cambridge, UK and Nancy  
 Shepherd, Dupont Merck, Wilmington, DE. Genomic DNA was  
 isolated from Trypanosoma brucei (stock TREU927/4) and  
 partially digested with Sau 3A1. DNA fragments were cloned  
 into the Bam HI site of pAD10SacBII vector (Genbank  
 accession U09128). The average insert size is 65 Kb.  
 Coverage: approx 4.4 X the haploid non-minichromosomal  
 genome."

BASE COUNT 136 a 161 c 157 g 119 t

ORIGIN

Alignment Scores:  
 Pred. No.: 56.5 Length: 573  
 Score: 54.00 Matches: 10  
 Percent Similarity: 84.62% Conservative: 1  
 Best Local Similarity: 76.92% Mismatches: 2  
 Query Match: 60.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-771-009-1 (1-17) x AQ639839 (1-573)

Qy 5 LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 A2194671  
 DB 71 TTGCTGCTGTCGGGTGATGTGAAGAGAAATCCCGGCCCG 109  
 ||| |||:|||||:|||||:|||||:|||||

RESULT 8  
 A2194671  
 LOCUS SP\_1028\_A2\_G12\_T7A Strongylocentrotus purpuratus, purple sea urchin,  
 sperm genomic BAC library Strongylocentrotus purpuratus genomic  
 clone Plate-1028 Col=24 Row=M, DNA sequence.

ACCESSION A2194671  
 VERSION A2194671.1 GI:8377850  
 KEYWORDS GSS.

```

SOURCE
ORGANISM      Strongylocentrotus purpuratus.
               Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
               Echinoidea; Euechinozoa; Echinacea; Echinoida;
               Strongylocentrotidae; Strongylocentrotus.

REFERENCE
AUTHORS       Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
               Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
               ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and
               Hood,L.
TITLE         A sea urchin genome project: Sequence scan, virtual map, and
               additional resources
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
MEDLINE        20402566
COMMENT       Contact: Cameron, RA, Davidson, EH, Hood, L
               Division of Biology 156-29
               California Institute of Technology
               Pasadena California 91125, USA
               Tel: (626) 395-8421
               Fax: (626) 793-3047
               Email: acameron@caltech.edu
               Plate: 1028 row: M column: 24
               Seq primer: T7
               Class: BAC ends
               High quality sequence stop: 781.

FEATURES
source        Location/Qualifiers
               1..781
               /organism="Strongylocentrotus purpuratus"
               /db_xref="taxon:7668"
               /clone="plate=1028 Col=24 Row=M"
               /clone_lib="Strongylocentrotus purpuratus, purple sea
               urchin, sperm genomic BAC library"
               /notes="Organ: sperm; Vector: BACs3.6; BAC Clones in E-Coli
               DH10B"

BASE COUNT    229 a 173 c 132 g 247 t

Alignment Scores:
Pred. No.:    96.6      Length: 781
Score:        53.50     Matches: 12
Percent Similarity: 82.35% Conservative: 2
Best Local Similarity: 70.59% Mismatches: 2
Query Match:  59.44%   Indels: 1
DB:           17       Gaps: 1

US-09-771-009-1 (1-17) x AZ194671 (1-781)

QY 2 PheAspLeuLeuLysLeu---AlaGlyAspValGluSerAsnProGlyPro 17
||||:||||| ||| ::||| ||||| ||||| ||||| |||||
Db 234 TTTGACTGCTCTTGAATCTGTGATGAGCTCAACCGAGCCCT 284

RESULT 9
AZ302518/c AZ302518 288 bp DNA linear GSS 09-JAN-2001
LOCUS       GSSTC12396 Trypanosoma cruzi random genomic library Trypanosoma
DEFINITION  cruzi genomic clone 657N10, DNA sequence.
ACCESSION   AZ302518
VERSION     AZ302518.1 GI:10128729
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi.
            Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1 (bases 1 to 288)
AUTHORS     Aguerro,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE       A random sequencing approach for the analysis of the trypanosoma
            cruzi genome: general structure, large gene and repetitive DNA
            families, and gene discovery
JOURNAL     Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE     20568489
COMMENT     Contact: Sanchez D.O.
            Instituto de Investigaciones Bioteconologicas (Univ. Nac. de Gral
            San Martin)

FEATURES
source        Location/Qualifiers
               1..326
               /organism="Trypanosoma cruzi"
               /strain="CL-Brener"

```

```

Av. Gral Paz entre Albarcellos y Constituyentes, INTI edificio 24
CP(1650) San Martin, Prov. de Bs AS. Argentina
Tel: 54-11-4580-7255 ext 309
Fax: 54-11-4752-9639
Email: dsanchez@iib.unsam.edu.ar
Sequences were basecalled with phred and vector was masked with
crossmatch (see http://genome.washington.edu). Sequences were then
trimmed from both ends to remove low quality bases and masked
vector.
Seq primer: T7
Class: shotgun.
Location/Qualifiers
1..288
/organism="Trypanosoma cruzi"
/strain="CL-Brener"
/db_xref="taxon:5693"
/clone="G57N10"
/clone_lib="Trypanosoma cruzi random genomic library"
/cell_type="epimastigote"
/notes="Vector: pBS(-) (Stratagene); T. cruzi DNA was
randomly sheared using a nebulizer and the 1 to 2 Kb range
was gel purified and cloned into the dephosphoryated
HincII site of the vector"
HincII site of the vector"
53 a 76 c 81 g 78 t

BASE COUNT    53 a 76 c 81 g 78 t
ORIGIN

Alignment Scores:
Pred. No.:    37.7      Length: 288
Score:        35.00     Matches: 9
Percent Similarity: 84.62% Conservative: 2
Best Local Similarity: 69.23% Mismatches: 2
Query Match:  58.89%   Indels: 0
DB:           17       Gaps: 0

US-09-771-009-1 (1-17) x AZ302518 (1-288)

QY 5 LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
||||:||||| ||| ::||| ||||| ||||| ||||| |||||
Db 124 CTACTGCTAAGCGGAGACATTGAACAGACAGCCAGGCCCC 86

RESULT 10
AQ904204 AQ904204 326 bp DNA linear GSS 09-JAN-2001
LOCUS       GSSTC09943 Trypanosoma cruzi random genomic library Trypanosoma
DEFINITION  cruzi genomic clone G4I13, DNA sequence.
ACCESSION   AQ904204
VERSION     AQ904204.1 GI:6484534
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi.
            Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1 (bases 1 to 326)
AUTHORS     Aguerro,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE       A random sequencing approach for the analysis of the trypanosoma
            cruzi genome: general structure, large gene and repetitive DNA
            families, and gene discovery
JOURNAL     Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE     20568489
COMMENT     Contact: Sanchez D.O.
            Instituto de Investigaciones Bioteconologicas (Univ. Nac. de Gral
            San Martin)

FEATURES
source        Location/Qualifiers
               1..326
               /organism="Trypanosoma cruzi"
               /strain="CL-Brener"

```

Best Local Similarity:	69.23%	Mismatches:	2
Query Match:	58.89%	Indels:	0
DB:	17	Gaps:	0
US-09-771-009-1 (1-17) x AQ908490 (1-326)			
Qy	5	LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro	17
		:::     :::	
Db	89	CTACTGCTAAGCGGAGACATTGAGCAGAAAGCCAGGCC	127

RESULT 12	AZ302155	359 bp	DNA	linear	GSS 09-JAN-2001
AZ302155/C	GSSTC12369	Trypanosoma cruzi	random genomic library		
LOCUS	AZ302155	359 bp	DNA	linear	GSS 09-JAN-2001
DEFINITION	GSSTC12369	Trypanosoma cruzi	random genomic library		
	AZ302155	359 bp	DNA	linear	GSS 09-JAN-2001
ACCESSION	AZ302155	359 bp	DNA	linear	GSS 09-JAN-2001
VERSION	AZ302155.1	359 bp	DNA	linear	GSS 09-JAN-2001
KEYWORDS	GSS.				

**SOURCE** Trypanosoma cruzi.  
**ORGANISM** Trypanosoma cruzi

**REFERENCE**  
**AUTHORS**  
 Acuña, F., Verdun, R., Frasch, A. C. C., and Sanchez, D. O.  
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma; Schizotrypanum.  
 1 (bases 1 to 359)

TITLE	A random sequencing approach for the analysis of the trypanosoma cruzi genome: general structure, large gene and repetitive DNA families, and gene discovery
JOURNAL	Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE	20568489
COMMENT	Contact: Sanchez D. O.

Instituto de Investigaciones Biotecnológicas (Univ. Nac. de Gral San Martín)  
Av. Gral Paz entre Albarillos y Constituyentes, INTI edificio 24  
CP(1650) San Martín, prov. de BS AS, Argentina  
Tel: 54-11-4580-7255 ext 309

Fax: 54-11-4752-9639  
 Email: dsanchez@iib.unsam.edu.ar

Sequences were basecalled with phred and vector was masked with crossmatch (see <http://genome.washington.edu>). Sequences were then trimmed from both ends to remove low quality bases and masked vector.

Seq primer: T7

Class: shotgun.

FEATURES	Location/Qualifiers
source	1 359

```

/organism="Trypanosoma cruzi"
/strain="CL-Brener"
/db_xref="taxon:5693"
/clone="G1608"
/clone_lib="Trypanosoma cruzi random genomic library"
/cell_type="epimastigote"
/notes="Vector: pBS(-) (Stratagene); T. cruzi DNA was randomly sheared using a nebulizer and the 1 to 2 Kb range was gel purified and cloned into the dephosphorvated

```

BASE COUNT  
ORIGIN

Pred. No.:	48.4	Length:	359
Score:	53.00	Matches:	9
Percent Similarity:	84.62%	Conservative:	2
Best Local Similarity:	69.23%	Mismatches:	2
Query Match:	58.89%	Indels:	0
DB:	17	Gaps:	0

US-09-771-009-1 (1-17) x AZ302155 (1-359)

Qy	5	LeuLysLeuLaGlyAspValGluSerAsnProGlyPro	17
		:::: : : : : : : : : : : : : : : : :	
Db	122	CTACTGTACGCGAGACATTGACAGACCCAGGCC	84

RESULT 13

```

A0445296/c
LOCUS           A0445296               379 bp    DNA          linear    GSS 09-JAN-2001
DEFINITION     GSSTc01571 Trypanosoma cruzi random genomic library Trypanosoma
                cruzi genomic clone G4C3, DNA sequence.
ACCESSION      A0445296
VERSION        A0445296.3  GI:10138167
KEYWORDS       GSS.
SOURCE         Trypanosoma cruzi.
ORGANISM       Trypanosoma cruzi
                Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
                Trypanosoma; Schizotrypanum.
REFERENCE      1 (bases 1 to 379)
AUTHORS        Aguero,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE          A random sequencing approach for the analysis of the trypanosoma
                cruzi genome: general structure, large gene and repetitive DNA
                families, and gene discovery
JOURNAL        Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE        20568489
COMMENT        On Sep 14, 2000 this sequence version replaced gi:9375285.
                Contact: Sanchez D.O.
                Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
                San Martin)
                Av. Gral Paz entre Albarcellos Y Constituyentes, INTI edificio 24
                CP(1650) San Martin, Prov. de BS AS. Argentina
                Tel: 54-11-4580-7255 ext 309
                Fax: 54-11-4752-9639
                Email: dsanchez@iib.unsam.edu.ar
                Sequences were basecalled with phred and vector was masked with
                crossmatch (see http://genome.washington.edu). Sequences were then
                trimmed from both ends to remove low quality bases and masked
                vector.
                Seq primer: T7
                Class: shotgun.
FEATURES       Location/Qualifiers
                source          1..379
                                /organism="Trypanosoma cruzi"
                                /strain="CL-Brener"
                                /db_xref="taxon:5693"
                                /clone="G4C3"
                                /cell_type="epimastigote"
                                /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
                                randomly sheared using a nebulizer and the 1 to 2 Kb range
                                was gel purified and cloned into the dephosphoryated
                                HincII site of the vector"
BASE COUNT     76 a   97 c   117 g   89 t
ORIGIN
Alignment Scores:
Pred. No.:      51.5          Length:      379
Score:          53.00         Matches:      9
Percent Similarity: 84.62%    Conservative: 2
Best Local Similarity: 69.23%  Mismatches:    2
Query Match:    58.89%        Indels:       0
DB:             17           Gaps:         0

US-09-771-009-1 (1-17) x A0445296 (1-379)
QY    5  LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
      ||| |||:|||||:||||| ||| ||| ||| |||
Db    111 CTACTGCTAAGCGGAGACATTGAGCAGACCCAGGCCCC 73

RESULT 14
A0445296/c
LOCUS           A0445296               418 bp    DNA          linear    GSS 09-JAN-2001
DEFINITION     GSSTc03451 Trypanosoma cruzi random genomic library Trypanosoma
                cruzi genomic clone G42N24, DNA sequence.
ACCESSION      A0910615
VERSION        A0910615.3  GI:10136876
KEYWORDS       GSS.
SOURCE         Trypanosoma cruzi.
ORGANISM       Trypanosoma cruzi
                Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
                Trypanosoma; Schizotrypanum.
REFERENCE      1 (bases 1 to 427)
AUTHORS        Aguero,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE          A random sequencing approach for the analysis of the trypanosoma
                cruzi genome: general structure, large gene and repetitive DNA
                families, and gene discovery
JOURNAL        Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE        20568489
COMMENT        Contact: Sanchez D.O.
                Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
                San Martin)
                Av. Gral Paz entre Albarcellos Y Constituyentes, INTI edificio 24
                CP(1650) San Martin, Prov. de BS AS. Argentina
                Tel: 54-11-4580-7255 ext 309
                Fax: 54-11-4752-9639
                Email: dsanchez@iib.unsam.edu.ar
                Sequences were basecalled with phred and vector was masked with
                crossmatch (see http://genome.washington.edu). Sequences were then
                trimmed from both ends to remove low quality bases and masked
                vector.
                Seq primer: T7
                Class: shotgun.
FEATURES       Location/Qualifiers
                source          1..418
                                /organism="Trypanosoma cruzi"
                                /strain="CL-Brener"
                                /db_xref="taxon:5693"
                                /clone="G42N24"
                                /cell_type="epimastigote"
                                /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
                                randomly sheared using a nebulizer and the 1 to 2 Kb range
                                was gel purified and cloned into the dephosphoryated
                                HincII site of the vector"
BASE COUNT     108 a   117 c   116 g   77 t
ORIGIN
Alignment Scores:
Pred. No.:      57.5          Length:      418
Score:          53.00         Matches:      9
Percent Similarity: 84.62%    Conservative: 2
Best Local Similarity: 69.23%  Mismatches:    2
Query Match:    58.89%        Indels:       0
DB:             17           Gaps:         0

US-09-771-009-1 (1-17) x A0910615 (1-418)
QY    5  LeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
      ||| |||:|||||:||||| ||| ||| ||| |||
Db    209 CTACTGCTAAGCGGAGACATTGAGCAGACCCAGGCCCC 247

RESULT 15
A0910615/c
LOCUS           A0910615               427 bp    DNA          linear    GSS 09-JAN-2001
DEFINITION     GSSTc12233 Trypanosoma cruzi random genomic library Trypanosoma
                cruzi genomic clone G18C8, DNA sequence.
ACCESSION      AZ302187
VERSION        AZ302187.1  GI:10128398
KEYWORDS       GSS.
SOURCE         Trypanosoma cruzi.
ORGANISM       Trypanosoma cruzi
                Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
                Trypanosoma; Schizotrypanum.
REFERENCE      1 (bases 1 to 427)
AUTHORS        Aguero,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
TITLE          A random sequencing approach for the analysis of the trypanosoma
                cruzi genome: general structure, large gene and repetitive DNA
                families, and gene discovery
JOURNAL        Genome Res. 10 (12), 1996-2005 (2000)
MEDLINE        20568489
COMMENT        Contact: Sanchez D.O.
                Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
                San Martin)
                Av. Gral Paz entre Albarcellos Y Constituyentes, INTI edificio 24
                CP(1650) San Martin, Prov. de BS AS. Argentina
                Tel: 54-11-4580-7255 ext 309
                Fax: 54-11-4752-9639
                Email: dsanchez@iib.unsam.edu.ar
                Sequences were basecalled with phred and vector was masked with
                crossmatch (see http://genome.washington.edu). Sequences were then
                trimmed from both ends to remove low quality bases and masked
                vector.
                Seq primer: T7
                Class: shotgun.
FEATURES       Location/Qualifiers
                source          1..418
                                /organism="Trypanosoma cruzi"
                                /strain="CL-Brener"
                                /db_xref="taxon:5693"
                                /clone="G42N24"
                                /cell_type="epimastigote"
                                /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
                                randomly sheared using a nebulizer and the 1 to 2 Kb range
                                was gel purified and cloned into the dephosphoryated
                                HincII site of the vector"
BASE COUNT     108 a   117 c   116 g   77 t
ORIGIN

```

San Martin)  
 Av. Gral Paz entre Albarellos y Constituyentes, INTI edificio 24  
 CP(1650) San Martin, Prov. de BS AS, Argentina  
 Tel: 54-11-4580-7255 ext 309  
 Fax: 54-11-4752-9639  
 Email: dsanchez@lib.unsam.edu.ar  
 Sequences were basecalled with phred and vector was masked with  
 crossmatch (see <http://genome.washington.edu>). Sequences were then  
 trimmed from both ends to remove low quality bases and masked  
 vector.

Seq primer: T7  
 Seq primer: T7  
 Class: shotgun.

FEATURES

Source

Location/Qualifiers

1..427  
 /organism="Trypanosoma cruzi"  
 /strain="CL-Brener"  
 /db\_xref="taxon:5693"  
 /clone="G18C8"  
 /clone\_lib="Trypanosoma cruzi random genomic library"  
 /cell\_type="epimastigote"  
 /note="Vector: PBS(-) (Stratagene); T. cruzi DNA was  
 randomly sheared using a nebulizer and the 1 to 2 kb range  
 was gel purified and cloned into the dephosphoryated  
 HincII site of the vector"

BASE COUNT 77 a 111 c 116 g 123 t  
 ORIGIN

Alignment Scores:

Pred. No.:	58.9	Length:	427
Score:	53.00	Matches:	9
Percent Similarity:	84.62%	Conservative:	2
Best Local Similarity:	69.23%	Mismatches:	0
Query Match:	58.89%	Indels:	0
DB:	17	Gaps:	0

US-09-771-009-1 (1-17) x AZ302187 (1-427)

QY 5 LeuLysIeuAlaGlyAspValGluSerAsnProGlyPro 17

Db 113 CTAGTCTAAGCGGAGACATTGAGCAGACCCAGGCCCC 75

Search completed: January 24, 2003, 20:16:20  
 Job time : 2236 secs





US-08-844-045C-16  
; Sequence 16, Application US/08844045C  
; Patent No. 6232099  
; GENERAL INFORMATION:  
; APPLICANT: Scottish Crop Research Institute  
; TITLE OF INVENTION: Method of Producing Chimeric Protein  
; FILE REFERENCE: Method of Producing a Chimeric Protein  
; CURRENT APPLICATION NUMBER: US/08/844,045C  
; CURRENT FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Foot-and-mouth disease virus  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: NIL  
; NAME/KEY: CDS  
; LOCATION: (1)..(60)  
US-08-844-045C-16

Alignment Scores:  
Pred. No.: 9.79e-08 Length: 60  
Score: 83.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.22% Indels: 0  
DB: 4 Gaps: 0

US-09-771-009-1 (1-17) x US-08-844-045C-16 (1-60)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProgly 16  
|||||  
Db 13 AATTTGACCTTCTCAAGTTGGCGGAGACGTCGAGTCCAACCCCTGGG 60

RESULT 3

US-08-844-045C-12  
; Sequence 12, Application US/08844045C  
; Patent No. 6232099  
; GENERAL INFORMATION:  
; APPLICANT: Scottish Crop Research Institute  
; TITLE OF INVENTION: Method of Producing Chimeric Protein  
; FILE REFERENCE: Method of Producing a Chimeric Protein  
; CURRENT APPLICATION NUMBER: US/08/844,045C  
; CURRENT FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 81  
; TYPE: DNA  
; ORGANISM: Foot-and-mouth disease virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(81)  
US-08-844-045C-12

Alignment Scores:  
Pred. No.: 1.41e-07 Length: 81  
Score: 83.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.22% Indels: 0  
DB: 4 Gaps: 0

US-09-771-009-1 (1-17) x US-08-844-045C-12 (1-81)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProgly 16  
|||||  
Db 34 AATTTGACCTTCTTAAGCTTGGCGGAGACGTCGAGTCCAACCCCTGGG 81

RESULT 4

US-08-844-045C-18  
; Sequence 18, Application US/08844045C

; Patent No. 6232099  
; GENERAL INFORMATION:  
; APPLICANT: Scottish Crop Research Institute  
; TITLE OF INVENTION: Method of Producing Chimeric Protein  
; FILE REFERENCE: Method of Producing a Chimeric Protein  
; CURRENT APPLICATION NUMBER: US/08/844,045C  
; CURRENT FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 183  
; TYPE: DNA  
; ORGANISM: Foot-and-mouth disease virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(183)  
US-08-844-045C-18

Alignment Scores:  
Pred. No.: 3.81e-07 Length: 183  
Score: 83.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.22% Indels: 0  
DB: 4 Gaps: 0

US-09-771-009-1 (1-17) x US-08-844-045C-18 (1-183)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProgly 16  
|||||  
Db 136 AATTTGACCTTCTCAAGTTGGCGGAGACGTCGAGTCCAACCCCTGGG 183

RESULT 5

US-09-091-219-1  
; Sequence 1, Application US/09091219  
; Patent No. 6171592  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/091,219  
; CURRENT FILING DATE: 1998-10-05  
; EARLIER APPLICATION NUMBER: PCT/AU96/00815  
; EARLIER FILING DATE: 1996-12-18  
; EARLIER APPLICATION NUMBER: AU PN7201  
; EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 7278  
; TYPE: DNA  
; ORGANISM: equine rhinovirus 1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (436)..(7176)  
US-09-091-219-1

Alignment Scores:  
Pred. No.: 7.84e-05 Length: 7278  
Score: 81.00 Matches: 15  
Percent Similarity: 94.12% Conservative: 1  
Best Local Similarity: 88.24% Mismatches: 1  
Query Match: 90.00% Indels: 0  
DB: 4 Gaps: 0

US-09-771-009-1 (1-17) x US-09-091-219-1 (1-7278)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProglyPro 17  
|||||  
Db 3409 AATTACTCTCTCTCAAAATTTGGCTGGAGATGTTGAGAGCAACCCCTGGCCCC 3459



```
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: genomic DNA
/ ; HYPOTHETICAL: NO
/ ; PUBLICATION INFORMATION:
/ ; AUTHORS: PEOPLES
/ ; JOURNAL: J. Biol. Chem.
/ ; VOLUME: 264
/ ; PAGES: 15298-15303
/ ; DATE: 1989
/ ;
US-08-241-943-25
Alignment Scores:
Pred. No.: 192 Length: 1770
Score: 42.00 Matches: 9
Percent Similarity: 62.50% Conservative: 1
Best Local Similarity: 56.25% Mismatches: 6
Query Match: 46.67% Indels: 0
DB: 1 Gaps: 0

US-09-771-009-1 (1-17) x US-08-241-943-25 (1-1770)
QY 2 PheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
|||||
Db 1258 TTGGACCTGCTGTCTGTGGAACGGCGACGCCACCACTGCGGGGCG 1305

RESULT 9
US-08-254-357-3
; Sequence 3, Application US/08254357
; Patent No. 5610041
; GENERAL INFORMATION:
; APPLICANT: Christopher R. Somerville,
; APPLICANT: Christiane Nawrath,
; APPLICANT: Yves Poirier
; TITLE OF INVENTION: Processes For Producing
; TITLE OF INVENTION: Polyhydroxybutyrate and Related
; TITLE OF INVENTION: Polyhydroxyalkanoates in the
; TITLE OF INVENTION: Plastids of Higher Plants
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: Acer
; OPERATING SYSTEM: MS-DOS (version 3.3)
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,357
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/108,193 and 07/732,243
; FILING DATE: August 17, 1993 and July 19, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-222
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2019 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE:
```

```
/ ; DESCRIPTION: Genomic DNA
/ ; HYPOTHETICAL: NO
/ ; ANTI-SENSE: NO
/ ; ORIGINAL SOURCE:
/ ; ORGANISM: Alcaligenes eutrophus
/ ; IMMEDIATE SOURCE:
/ ; LIBRARY: Genomic
/ ;
US-08-254-357-3
Alignment Scores:
Pred. No.: 225 Length: 2019
Score: 42.00 Matches: 9
Percent Similarity: 62.50% Conservative: 1
Best Local Similarity: 56.25% Mismatches: 6
Query Match: 46.67% Indels: 0
DB: 1 Gaps: 0

US-09-771-009-1 (1-17) x US-08-254-357-3 (1-2019)
QY 2 PheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
|||||
Db 1507 TTGGACCTGCTGTCTGTGGAACGGCGACGCCACCACTGCGGGGCG 1554

RESULT 10
US-08-778-570B-8
; Sequence 8, Application US/08778570B
; Patent No. 6437096
; GENERAL INFORMATION:
; APPLICANT: Myers, Lisa E
; APPLICANT: Schryvers, Anthony B
; APPLICANT: Harkness, Robin E
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Du, Run-Pan
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: Transferrin Receptor Genes of Moraxella
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,570B
; FILING DATE: 03-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-664
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2127 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-778-570B-8
Alignment Scores:
Pred. No.: 240 Length: 2127
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
```



```
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: Transferrin Receptor Genes of Moraxella
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,570B
; FILING DATE: 03-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-564
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3435 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-778-570B-7

Alignment Scores:
Pred. No.: 429 Length: 3435
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 46.67% Indels: 0
DB: 4 Gaps: 0

US-09-771-009-1 (1-17) x US-08-778-570B-7 (1-3435)
QY 3 AspLeuLeuLysLeuAlaGlyAspValGluSerAsnPro 15
|||||:||||| |||||:|||||
Db 1083 GATGTGTTAAACTGGAAGTGACTTGAAGCATAATCCA 1121

RESULT 14
US-09-059-584-7
; Sequence 7, Application US/09059584
; Patent No. 6440701
; GENERAL INFORMATION:
; APPLICANT: Myers, Lisa E
; APPLICANT: Schryvers, Anthony B
; APPLICANT: Harkness, Robin E
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Du, Run-Pan
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: Transferrin Receptor Genes of Moraxella
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,584
; FILING DATE: 14-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/778,570
; FILING DATE: 03-JAN-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-794
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3435 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-059-584-7

Alignment Scores:
Pred. No.: 429 Length: 3435
Score: 42.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 46.67% Indels: 0
DB: 4 Gaps: 0

US-09-771-009-1 (1-17) x US-09-059-584-7 (1-3435)
QY 3 AspLeuLeuLysLeuAlaGlyAspValGluSerAsnPro 15
|||||:||||| |||||:|||||
Db 1083 GATGTGTTAAACTGGAAGTGACTTGAAGCATAATCCA 1121

RESULT 15
US-08-472-358-1
; Sequence 1, Application US/08472358
; Patent No. 5650555
; GENERAL INFORMATION:
; APPLICANT: Chris Somerville, Yves Poirier,
; APPLICANT: Douglas Dennis
; TITLE OF INVENTION: Transgenic Plant Materials
; TITLE OF INVENTION: Producing Polyhydroxyalkanoates
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM AT
; OPERATING SYSTEM: MS-DOS (version 3.3)
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,358
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/732,243
; FILING DATE: July 19, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-131
; TELECOMMUNICATION INFORMATION:
```

```

; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4983 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Alcaligenes eutrophus
; IMMEDIATE SOURCE:
; LIBRARY: Genomic
; US-08-472-358-1

Alignment Scores:
Pred. No.: 675 Length: 4983
Score: 42.00 Matches: 9
Percent Similarity: 62.50% Conservative: 1
Best Local Similarity: 56.25% Mismatches: 6
Query Match: 46.67% Indels: 0
DB: 1 Gaps: 0

US-09-771-009-1 (1-17) x US-08-472-358-1 (1-4983)
QY 2 PheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
Db 2102 TTGACCTGCTGTCTGTGGAACGGCGACGCCACCACTGCCGGGGCCG 2149

```

Search completed: January 24, 2003, 20:17:40  
Job time : 70 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 24, 2003, 18:38:14 ; Search time 301 Seconds  
(without alignments)  
127.189 Million cell updates/sec

Title: US-09-771-009-1

Perfect score: 90

Sequence: 1 NFDLLKLAGDVESNPGP 17

Scoring table:  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool/US09771009/runat\_22012003\_140100\_18416/app\_query.fasta\_1.199  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOCPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09771009.ecgn.1.1.79.0runat\_22012003\_140100\_18416 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	72	21	AAA99502
2	90	100.0	488	21	AAZ99338
3	90	100.0	488	21	AAZ99338
4	90	100.0	575	21	AAZ99339
5	90	100.0	873	24	ABK31150
6	90	100.0	939	24	AAZ35101
7	90	100.0	961	2	AA10025
8	90	100.0	1105	24	ABK31034
9	90	100.0	5715	21	AAZ34936
10	83	92.2	60	18	AAZ30945
11	83	92.2	60	18	AAZ2989
12	83	92.2	81	18	AAZ2990
13	83	92.2	183	18	AAZ2988
14	81	90.0	7277	18	AAZ2991
15	80	88.9	2802	3	AAZ85178
16	80	88.9	3444	24	AAZ20016
17	70	77.8	648	5	ABL53056
18	70	77.8	702	5	AAZ40074
19	70	77.8	711	5	AAZ40075
20	48	53.3	2535	22	AAZ77900
21	47	52.2	1950	23	ABL08841
22	47	52.2	3219	23	ABL08079
23	47	52.2	4043	23	ABL08840
24	47	52.2	15339	23	ABL08078
25	47	52.2	25003	22	AAZ19679
26	46	51.1	356	22	AAZ60725
27	46	51.1	407	22	AAZ78397
28	46	51.1	417	24	ABN26099
29	46	51.1	870	23	AAZ9520
30	46	51.1	1619	23	AAZ30308
31	46	51.1	1767	23	AAZ64805
32	46	51.1	1767	23	AAZ75105
33	46	51.1	19468	23	ABL06926
34	46	51.1	21399	23	ABL21134
35	45	50.0	654	21	AAZ12751
36	45	50.0	1152	21	AAZ20015
37	45	50.0	49380	23	ABL11838
38	44	48.9	243	24	ABN66750
39	44	48.9	413	21	AAZ36055
40	44	48.9	755	21	AAH51423
41	44	48.9	991	21	AAH51424
42	44	48.9	2733	22	AAZ61028
43	44	48.9	17245	22	AAZ83897
44	44	48.9	31882	23	ABL09072
45	44	48.9	2365589	24	ABA90521

ALIGNMENTS

RESULT 1  
AAA99502  
ID AAA99502 standard; DNA; 72 BP.  
XX AAA99502;  
AC AAA99502;  
XX  
DT 15-FEB-2001 (first entry)  
XX  
DE PRRSV attenuated virus vaccine PCR primer LV139(14609).  
XX  
KW PRRSV; Lelystad virus; pig viral infection; vaccine; attenuated virus;  
KW PCR primer; ss.  
XX  
OS Porcine reproductive and respiratory syndrome virus.  
XX  
PN WO200053787-A1.  
XX  
PD 14-SEP-2000.

XX 08-MAR-2000; 2000WO-NL00152.  
 XX  
 XX 08-MAR-1999; -99EP-0200668.  
 XX  
 XX (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.  
 XX  
 XX Meulenber J, Van Nieuwstadt AP, Langeveld J;  
 XX  
 XX WPI; 2000-594332/56.  
 XX

XX Porcine reproductive and respiratory syndrome virus replicon useful as  
 PT a non-spreading and marker vaccine, has deletions of certain nucleic  
 PT acid of original virus and is capable of in vivo RNA replication  
 XX  
 XX Disclosure; Page 40; 52pp; English.  
 XX

XX The present sequence is one of the PCR primers used during the production  
 CC of a vaccine against the porcine reproductive and respiratory syndrome  
 CC virus (PRRSV, also known as Lelystad virus). This virus causes  
 CC respiratory problems in pigs and abortions in sows. The N-protein can be  
 CC mutated to produce an attenuated virus which acts as a vaccine in pigs  
 CC against PRRSV. This has the advantage that it is highly immunogenic with  
 CC less virulence than previous vaccines.  
 XX

SQ Sequence 72 BP; 20 A; 22 C; 19 G; 11 T; 0 other;

Alignment Scores:  
 Pred. No.: 4.87e-08 Length: 72  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-771-009-1 (1-17) x AAA99502 (1-72)

OY 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 DB 1 AACTTTGACCTTCTCAAGTTGGCCGCGACGTCGAGTCCACCCAGGGCCC 51

RESULT 2  
 AAZ99338  
 ID AAZ99338 standard; DNA; 488 BP.  
 AC AAZ99338;  
 XX  
 XX 03-JUL-2000 (first entry)  
 DT  
 XX DNA encoding a fusion protein of DmAMP1 and RsAFP2.  
 DE  
 XX Antimicrobial protein; AMP1; transgenic plant; linker propeptide;  
 KW protein expression; plant defensin; RsAFP2; antifungal protein; AFP2; ss.  
 XX  
 XX Synthetic.  
 OS  
 OS Dahlia merckii.  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 FH 3..479  
 CDS /\*tag= a  
 FT /product= "fusion protein of DmAMP1 and RsAFP2"  
 FT  
 XX

WO200011175-A1.  
 XX  
 XX 02-MAR-2000.  
 XX  
 XX 17-AUG-1999; 99WO-GB02716.  
 XX  
 XX 18-AUG-1998; 98GB-0018001.  
 PR  
 XX 04-DEC-1998; 98GB-0026753.  
 XX  
 XX (ZENE ) ZENECA LTD.  
 PA

XX Broekaert WF, Francois IEJA, De Bolle MFC, Evans IJ, Ray JA;  
 XX  
 XX WPI; 2000-246564/21.  
 DR P-PSDB; AAY84071.  
 XX  
 XX Improving expression of polyproteins in plants involves coexpression of  
 PT two or more proteins in plants within a single transcription unit  
 PT  
 XX Disclosure; Fig 33; 151pp; English.  
 XX  
 XX The present sequence encodes a protein of the invention, comprising  
 CC the mature proteins of the plant defensins, the Dahlia antimicrobial  
 CC protein (AMP) 1 and the antifungal protein 2 (RsAFP2), linked by a linker  
 CC propeptide of the invention. The specification describes methods for  
 CC improving expression levels of one or more proteins in a transgenic  
 CC plant. The method comprises inserting a DNA sequence having a promoter  
 CC region operably linked to two or more protein encoding regions separated  
 CC by a DNA sequence coding for a linker propeptide and a terminator region.  
 CC The method is used to produce proteins in plants. The linker propeptide  
 CC comprising a cleavage site, whereby the expressed polypeptide is  
 CC post-translationally processed into the component protein molecules.  
 CC The propeptide sequence is rich in amino acids A, V, S and T and  
 CC contains dipeptidic sequences consisting of either two acidic, two  
 CC basic or one acidic and one basic residue as a cleavable linker  
 CC sequence.  
 XX

SQ Sequence 488 BP; 117 A; 116 C; 131 G; 124 T; 0 other;

Alignment Scores:  
 Pred. No.: 4.79e-07 Length: 488  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-771-009-1 (1-17) x AAZ99338 (1-488)

OY 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 DB 273 AATTTGACCTTCTTAAGCTTGGGGAGACGTCGAGTCCAAACCTGGGCC 323

RESULT 3  
 AAZ99339  
 ID AAZ99339 standard; DNA; 575 BP.  
 AC AAZ99339;  
 XX  
 XX 03-JUL-2000 (first entry)  
 DT  
 XX DNA encoding a fusion protein of DmAMP1 and RsAFP2.  
 DE  
 XX Antimicrobial protein; AMP1; transgenic plant; linker propeptide;  
 KW protein expression; plant defensin; RsAFP2; antifungal protein; AFP2; ss.  
 XX  
 XX Synthetic.  
 OS  
 OS Dahlia merckii.  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 FH 3..566  
 CDS /\*tag= a  
 FT /product= "fusion protein of DmAMP1 and RsAFP2"  
 FT  
 XX  
 XX WO200011175-A1.  
 XX  
 XX 02-MAR-2000.  
 XX  
 XX 17-AUG-1999; 99WO-GB02716.  
 XX  
 XX 18-AUG-1998; 98GB-0018001.  
 PR  
 XX 04-DEC-1998; 98GB-0026753.  
 PR

XX (ZENE ) ZENECA LTD.  
XX Broekaert WF, Francois IEJA, De Bolle MFC, Evans IJ, Ray JA;  
PI WPI: 2000-246564/21.  
XX P-PSDB; AA784072.  
XX  
PT Improving expression of polyproteins in plants involves coexpression of  
PT two or more proteins in plants within a single transcription unit  
XX  
XX Disclosure; Fig 34; 151pp; English.  
XX  
XX The present sequence encodes a protein of the invention, comprising  
CC the mature proteins of the plant defensins, the Dahlia antimicrobial  
CC protein (AMP) 1 and the antifungal protein 2 (RsAFP2), linked by a linker  
CC propeptide of the invention. The specification describes methods for  
CC improving expression levels of one or more proteins in a transgenic  
CC plant. The method comprises inserting a DNA sequence having a promoter  
CC region operably linked to two or more protein encoding regions separated  
CC by a DNA sequence coding for a linker propeptide and a terminator region.  
CC The method is used to produce proteins in plants. The linker propeptide  
CC comprising a cleavage site, whereby the expressed polypeptide is  
CC post-translationally processed into the component protein molecules.  
CC The propeptide sequence is rich in amino acids A, V, S and T and  
CC contains dipeptidic sequences consisting of either two acidic, two  
CC basic or one acidic and one basic residue as a cleavable linker  
XX  
XX Sequence 575 BP; 133 A; 137 C; 149 G; 156 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 5.82e-07 Length: 575  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0  
US-09-771-009-1 (1-17) x AA299339 (1-575)  
QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
Db 273 AATTTTGACCTTCTTAAGCTTGGGAGAGCTGAGTCCAAACCTGGGCC 323  
RESULT 4  
ABK31150  
ID ABK31150 standard; cDNA; 873 BP.  
XX AC ABK31150;  
XX  
XX 23-APR-2002 (first entry)  
XX  
XX Plant dwarfing/stunting related cDNA seq ID 564.  
DE  
XX  
XX Plant; ss; dwarfism; stunting; EST; expressed sequence tag;  
KW transgenic plant; plant metabolism.  
XX  
XX Planta.  
OS  
XX WO200208410-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 20-JUL-2001; 2001WO-US23120.  
XX  
XX 20-JUL-2000; 2000US-219809P.  
XX 20-JUL-2000; 2000US-219810P.  
XX  
XX (DOWC ) DOW CHEM CO.  
PA (REDD/) REDDY S A.  
PA (LARR/) LARRINUA M I.  
PA (RUEG/) RUEGGER M.

PA (WEGL/) WEGLARZ T.  
PA (BLAK/) BLAKESLEE B.  
PA (ORIE/) ORIEDO V B J.  
PA (SAVI/) SAVICKAS J P.  
PA (MCCR/) MCCREY A D.  
PA (MILL/) MILLER A B.  
PA (GACH/) GACHOTTE D.  
PA (GROS/) GROSLEY R.  
PA (PELL/) PELL R.  
XX  
XX Reddy SA, Larrinua MI, Ruegger M, Weglarz T, Blakeslee B;  
PI Oriedo VB, Savickas JP, McGreery AD, Miller AB, Pogue PG;  
PI Della-Cioppa RG, Wolfe MG, Zheng W, Gachotte D, Grosley R, Pell R;  
XX WPI: 2002-164823/21.  
XX  
XX Polynucleotide and amino acid sequences identified in one or more  
PT metabolic pathways that lead to dwarfism and stunting in plants, useful  
PT in agriculture to create dwarf varieties of any plant species -  
XX  
XX Disclosure; Fig 9; 717pp; English.  
XX  
XX The invention relates to polynucleotide and amino acid sequences  
CC identified in one or more metabolic pathways that lead to dwarfism and  
CC stunting in plants. Also included are vectors comprising the  
CC polynucleotides, transgenic plants (including the seed and leaf)  
CC transfected with the polynucleotides or vectors, a process for altering  
CC the metabolism of a plant comprising providing the above vector and a  
CC plant, and transfecting the plant with the vector under conditions such  
CC that the metabolism of the plant is altered by expression of the isolated  
CC nucleic acid from the vector, e.g. such that a stunting phenotype in an  
CC industrial plant is produced, and a process for the characterisation of  
CC fractionated biological samples, comprising (a) providing one or more  
CC fractionated biological samples, references samples, a gas chromatography  
CC apparatus, a mass spectroscopy apparatus or data analysis software and  
CC (b) treating the fractionated biological samples and the reference  
CC samples with the gas chromatography apparatus to generate chromatographic  
CC data corresponding to the fractionated biological samples and the  
CC reference samples, (c) treating the fractionated biological samples and  
CC the reference samples with the mass spectroscopy apparatus to generate  
CC spectroscopic data corresponding to the fractionated biological samples  
CC and the reference samples and (d) processing the chromatographic and the  
CC spectroscopic data with the data analysis software. The nucleic acid and  
CC the vector are useful for altering the metabolism of a plant and for  
CC stunting a plant. The nucleic acids are useful in agriculture to create  
CC dwarf varieties of any plant species. The present sequence is a  
CC plant cDNA contig or singleton (related to dwarfism/stunting) identified  
CC by searching a nucleic acid database with plant EST (expressed  
CC sequence tag) and a BLAST (basic local alignment tool) stringency  
CC of e-20.  
XX  
XX Sequence 873 BP; 201 A; 203 C; 239 G; 230 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 9.58e-07 Length: 873  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
US-09-771-009-1 (1-17) x ABK31150 (1-873)  
QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
Db 686 AATTTTGATTTGCTGAAGTTGGGGGTGACGTGGATCTAACCTGGTCT 736  
RESULT 5  
AAD35101  
ID AAD35101 standard; DNA; 939 BP.  
XX  
XX AAD35101;  
XX



PA (REDD/) REDDY S A.  
PA (LARE/) LARRINUA M I.  
PA (RUEG/) RUEGGER M.  
PA (WEG/) WEGLARZ T.  
PA (BLAK/) BLAKESLEE B.  
PA (ORIE/) ORIEDO V B J.  
PA (SAVI/) SAVICKAS J P.  
PA (MCCR/) MCCREARY A D.  
PA (MILL/) MILLER A B.  
PA (GACH/) GACHOTTE D.  
PA (GROS/) GROSLEY R.  
PA (PELL/) PELL R.

PI Reddy SA, Larrinua MI, Ruegger M, Weglarz T, Blakeslee B;  
PI Oriedo VB, Savickas JP, McCreary AD, Miller AB, Poque PG;  
PI Della-Cioppa RG, Wolfe MG, Zheng W, Gachotte D, Grosley R, Pell R;  
PI WPI; 2002-164823/21.

XX Polynucleotide and amino acid sequences identified in one or more  
PT metabolic pathways that lead to dwarfism and stunting in plants, useful  
PT in agriculture to create dwarf varieties of any plant species -  
XX  
PS Disclosure; Fig 9; 717pp; English.

CC The invention relates to polynucleotide and amino acid sequences  
CC identified in one or more metabolic pathways that lead to dwarfism and  
CC stunting in plants. Also included are vectors comprising the  
CC polynucleotides, transgenic plants (including the seed and leaf)  
CC transfected with the polynucleotides or vectors, a process for altering  
CC the metabolism of a plant comprising providing the above vector and a  
CC plant, and transfecting the plant with the vector under conditions such  
CC that the metabolism of the plant is altered by expression of the isolated  
CC nucleic acid from the vector, e.g. such that a stunting phenotype in an  
CC industrial plant is produced and a process for the characterisation of  
CC fractionated biological samples, comprising (a) providing one or more  
CC fractionated biological samples, references samples, a gas chromatography  
CC apparatus, a mass spectroscopy apparatus or data analysis software and  
CC (b) treating the fractionated biological samples and the reference  
CC samples with the gas chromatography apparatus to generate chromatographic  
CC data corresponding to the fractionated biological samples and the  
CC reference samples, (c) treating the fractionated biological samples and  
CC the reference samples with the mass spectroscopy apparatus to generate  
CC spectroscopic data corresponding to the fractionated biological samples  
CC and the reference samples and (d) processing the chromatographic and the  
CC spectroscopic data with the data analysis software. The nucleic acid and  
CC the vector are useful for altering the metabolism of a plant and for  
CC stunting a plant. The nucleic acids are useful in agriculture to create  
CC dwarf varieties of any plant species. The present sequence is a  
CC plant cDNA contig or singleton (related to dwarfism/stunting) identified  
CC by searching a nucleic acid database with plant EST (expressed  
CC sequence tag) and a BLAST (basic local alignment tool) stringency  
CC of e-20.

XX  
SQ Sequence 1105 BP; 308 A; 265 C; 262 G; 270 T; 0 other;

Alignment Scores:			
Pred. No.:	1-27e-06	Length:	1105
Score:	90.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-771-009-1 (1-17) x ABK31034 (1-1105)

QY 1 AspPheAspLeuLeuysLeuAlaglyAspValGluSerAsnProGlyPro 17  
|||||

Db 310 AATTTTCATTGCTGAAGTTGGCGGTGACGTGGAATCTAACCTGGTCTCT 260  
|||||

RESULT 8

AAZ34936

ID AAZ34936 standard; DNA; 5715 BP.

XX  
AC AAZ34936;  
XX  
DT 28-FEB-2000 (first entry)  
XX  
DE Retrovirus vector.  
XX  
KW Interleukin-4 inducible epsilon promoter; IgE; immunoglobulin;  
KW antibody; modulator; screening; human; allergy; therapy;  
KW retrovirus; vector; green fluorescent protein; GFP; ss.  
XX

OS Chimeric - cytomegalovirus.  
OS Chimeric - Aequoria victoria.  
OS Chimeric - unidentified.

FH Key Location/Qualifiers  
FT promoter 1..845  
FT /\*tag= a

FT mutation /note= "CMV promoter/R/US 5' LTR"  
FT replace(1322,G)  
FT /\*tag= b

FT FT /\*note= "GAG ATG-ATC mutation"  
FT CDS 2151..2865  
FT /\*tag= c

FT FT /product= "green fluorescent protein"  
FT misc\_feature 2895..2952  
FT /\*tag= d

FT LTR /note= "FMDV 2a cleavage sequence"  
FT 3052..3645  
FT /\*tag= e

FT FT /note= "3' LTR"  
FT misc\_feature 3652..5715  
FT /\*tag= f

FT FT /note= "pGEM backbone (pUC origin, ampr)"  
XX  
PN W09958663-A1.

XX

XX 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10497.

XX 12-MAY-1998; 98US-0076624.

XX (RIGE-) RIGEL PHARM INC.

XX Ferrick DA, Swift SE, Armstrong R, Fox B;

XX WPI; 2000-062297/05.

XX Methods, cell lines and vectors for screening for modulators of  
XX immunoglobulin E synthesis, secretion and switch rearrangement -  
XX Disclosure; Fig 11B; 81pp; English.

XX This is the nucleotide sequence of a retrovirus vector preferred  
CC for use in methods of the invention. The invention provides methods  
CC of screening for bioactive agents capable of inhibiting the human  
CC interleukin-4 (IL-4) inducible epsilon promoter (see AAZ34932). The  
CC method comprises combining a candidate bioactive agent with a cell  
CC comprising a fusion nucleic acid composed of the IL-4 inducible  
CC epsilon promoter and a reporter gene. The promoter is then induced  
CC with IL-4 or IL-13, and the presence or absence of the reporter  
CC gene is detected. Absence of the reporter gene indicates that the  
CC agent inhibits the promoter. Preferred embodiments use retroviral  
CC vectors to introduce the candidate bioactive agents. Also provided  
CC are methods of screening for bioactive agents capable of modulating  
CC IgE synthesis, secretion and switch rearrangement. These methods  
CC rely on reporter genes fused to IgE promoters, such as the IL-4  
CC inducible epsilon promoter that starts a cascade that ultimately  
CC results in IgE production. The methods screen for upstream  
CC modulators of IgE production to prevent the production of IgE and  
CC thus reduce or eliminate the allergic response.

SQ Sequence 5715 BP; 1342 A; 1610 C; 1419 G; 1344 T; 0 other;

Alignment Scores:  
 Pred. No.: 9.03e-06 Length: 5715  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-771-009-1 (1-17) x AAF30945 (1-5715)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 |||||||  
 Db 2904 AATTTTGACCTTCTTAAACTTGGGGGAGCGTCGAGTCCCAACCTGGGCC 2954

RESULT 9

AAF30945

ID AAF30945 standard; DNA; 5715 BP.

XX AAF30945;

XX AAF30945;

DT 23-JUL-2001 (first entry)

XX Vector used in invention.

XX Interleukin-4 inducible epsilon promoter; human; IgE; antibody;

KW immunoglobulin E; allergy; therapy; switch rearrangement; vector;

KW CMV; green fluorescent protein; ds.

XX Chimeric - Cytomegalovirus.

OS Chimeric - Unspecified.

XX Key Location/Qualifiers

FT LTR 1..845

FT /tag= a

FT /note= "5' LTR"

FT promoter 1..845

FT /tag= b

FT /note= "CMV promoter"

FT mutation replace(1322,G)

FT /tag= c

FT misc\_feature 850..2100

FT /tag= d

FT /note= "extended psi region"

FT CDS 2151..2865

FT /tag= e

FT /product= "green fluorescent protein fusion with

FT C-terminal GGGGGG linker and FMDV 2a

FT cleavage sequence"

FT LTR 3052..3645

FT /tag= f

FT /note= "3' LTR"

FT misc\_feature 3652..5715

FT /tag= g

FT /note= "pGEM backbone (pUC origin, ampr)"

XX WO200134806-A2.

XX 17-MAY-2001.

XX 13-NOV-2000; 2000WO-US31232.

XX 12-NOV-1999; 99US-0165189.

XX (RIGF-) RIGEL PHARM INC.

XX Kinsella TM;

XX WPI; 2001-335931/35.

XX Screening for agents capable of inhibiting a promoter, especially

PT interleukin-4 inducible epsilon promoter involved in immunoglobulin E

PT production, by using diphtheria toxin constructs -  
 XX Disclosure; Fig 11B-1-11B-3; 80pp; English.  
 XX The present sequence is that of a vector preferred for use in  
 CC methods of the invention. These methods utilise diphtheria toxin  
 CC for screening purposes, especially for identifying modulators of  
 CC IgE synthesis, secretion and switch rearrangement. A claimed  
 CC method of screening for bioactive agents capable of inhibiting the  
 CC IL-4 inducible epsilon promoter, which is involved in IgE  
 CC switching, comprises: combining a candidate bioactive agent and a  
 CC cell that does not endogenously express heparin-binding epidermal  
 CC growth factor-like growth factor (HBEGF) and which comprises a  
 CC fusion nucleic acid comprising the IL-4 inducible epsilon promoter  
 CC and a nucleic acid encoding HBEGF; inducing the promoter with IL-4;  
 CC adding diphtheria toxin to the cell; and determining whether the  
 CC cell is dead. Compositions comprising a test vector and a reporter  
 CC vector that includes a reporter gene such as green fluorescent  
 CC proteins are provided. Inhibitors of IgE synthesis can be  
 CC identified that prevent the production of IgE and reduce or  
 CC eliminate an allergic response.

XX Sequence 5715 BP; 1342 A; 1610 C; 1419 G; 1344 T; 0 other;

Alignment Scores:

Pred. No.: 9.03e-06 Length: 5715

Score: 90.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 22 Gaps: 0

US-09-771-009-1 (1-17) x AAF30945 (1-5715)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17

|||||

Db 2904 AATTTTGACCTTCTTAAACTTGGGGGAGCGTCGAGTCCCAACCTGGGCC 2954

RESULT 10

AAF92989

ID AAT92989 standard; DNA; 60 BP.

XX AAT92989;

XX 24-APR-1998 (first entry)

XX Fragment from LITMUS 39 based plasmid pLit.GFP-2A16H-CP.

XX Virus like particle; diagnostic reagent; antibiotic; therapy; VLP;

XX food supplement; vaccination; herbicide resistance; industrial enzyme;

XX tobacco mild green mosaic virus; TMGMV; PCR primer; LITMUS 39; ss.

XX Synthetic.

XX Potato virus X.

XX WO9739134-A1.

XX 23-OCT-1997.

XX 17-APR-1997; 97WO-CB01065.

XX 17-APR-1996; 96GB-0007899.

XX (SCCR-) SCOTTISH CROP RES INST.

XX Chapman SN, Wilson TMA;

XX WPI; 1997-526468/48.

XX P-PSDB; AAW33911.

XX Production of virus like particles - using a nucleic acid sequence

XX capable of assembly with a protein having a first viral portion and

XX second non-viral portion

XX Example 2; Fig 4; 33pp; English.  
 PS This is a fragment from the LITMUS 39 based plasmid pLit.GFP-2A16H-CP.  
 CC These plasmids contain the green fluorescent protein (GFP)-2A- potato  
 CC virus X coat protein (CP) gene fusions. This is used as a source for  
 CC GFP-2A gene fusion with tobacco mild green mosaic virus (TMGMV) coat  
 CC protein (CP) gene and 3' UTR. This is used in a novel method for  
 CC producing a protein having a first (viral) portion and a second  
 CC (non-viral) portion. The method comprises, expressing the protein in a  
 CC cell, providing a nucleic acid sequence capable of assembly with the  
 CC protein into a virus-like particle (VLP), and permitting in vivo assembly  
 CC of the protein and nucleic acid into VLPs. The VLPs can be used for the  
 CC production of proteins such as diagnostic reagents, antibiotics,  
 CC therapeutic agents or food supplements. They can be used for e.g.  
 CC expression of metabolic enzymes for pathway engineering, nutritional  
 CC supplements, anti-potato cyst nematode lectins, gut protease inhibitors,  
 CC anti-botrytis agents, PGIPs, anti-insect Bacillus thuringiensis toxin and  
 CC herbicide resistance agents, industrial enzymes, pharmaceuticals,  
 CC therapeutic proteins, and nucleic acids, and as bioreactors. They can  
 CC also be used intact for presentation of peptide epitopes for vaccination  
 CC of animals, the production of therapeutic or industrial proteins and  
 CC polypeptides and/or the delivery of therapeutic nucleic acid molecules.  
 XX  
 SQ Sequence 60 BP; 13 A; 15 C; 16 G; 16 T; 0 other;

Alignment Scores:  
 Pred. No.: 7.16e-07 Length: 60  
 Score: 83.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 92.22% Indels: 0  
 DB: 18 Gaps: 0  
 US-09-771-009-1 (1-17) x AAT92989 (1-60)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGly 16  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 13 AATTTTGACCTTCTTAAGCTTGGCGGAGACGTCGAGTCCAAACCTGGG 60

RESULT 11  
 AAT92990  
 ID AAT92990 standard; DNA; 60 BP.  
 XX  
 AC AAT92990;  
 XX  
 DT 24-APR-1998 (first entry)  
 XX  
 DE Fragment from LITMUS 39 based plasmid pLit.GFP-2A16K-CP.  
 XX  
 KW Virus like particle; diagnostic reagent; antibiotic; therapy; VLP;  
 KW food supplement; vaccination; herbicide resistance; industrial enzyme;  
 KW tobacco mild green mosaic virus; TMGMV; PCR primer; LITMUS 39; ss.  
 XX  
 OS Synthetic.  
 OS Potato virus X.  
 XX  
 PN W09739134-A1.  
 XX  
 PD 23-OCT-1997.  
 XX  
 PF 17-APR-1997; 97WO-GB01065.  
 XX  
 PR 17-APR-1996; 96GB-0007899.  
 XX  
 PA (SCCR-) SCOTTISH CROP RES INST.  
 XX  
 PI Chapman SN, Wilson TMA;  
 XX  
 DR WPI; 1997-526468/48.  
 DR P-PSDB; AAW33911.  
 XX  
 PT Production of virus like particles - using a nucleic acid sequence

PT capable of assembly with a protein having a first viral portion and  
 PT second non-viral portion  
 XX  
 PS Example 2; Fig 4; 33pp; English.  
 XX This is a fragment from the LITMUS 39 based plasmid pLit.GFP-2A16K-CP.  
 CC These plasmids contain the green fluorescent protein (GFP)-2A- potato  
 CC virus X coat protein (CP) gene fusions. This is used as a source for  
 CC GFP-2A gene fusion with tobacco mild green mosaic virus (TMGMV) coat  
 CC protein (CP) gene and 3' UTR. This is used in a novel method for  
 CC producing a protein having a first (viral) portion and a second  
 CC (non-viral) portion. The method comprises, expressing the protein in a  
 CC cell, providing a nucleic acid sequence capable of assembly with the  
 CC protein into a virus-like particle (VLP), and permitting in vivo assembly  
 CC of the protein and nucleic acid into VLPs. The VLPs can be used for the  
 CC production of proteins such as diagnostic reagents, antibiotics,  
 CC therapeutic agents or food supplements. They can be used for e.g.  
 CC expression of metabolic enzymes for pathway engineering, nutritional  
 CC supplements, anti-potato cyst nematode lectins, gut protease inhibitors,  
 CC anti-botrytis agents, PGIPs, anti-insect Bacillus thuringiensis toxin and  
 CC herbicide resistance agents, industrial enzymes, pharmaceuticals,  
 CC therapeutic proteins, and nucleic acids, and as bioreactors. They can  
 CC also be used intact for presentation of peptide epitopes for vaccination  
 CC of animals, the production of therapeutic or industrial proteins and  
 CC polypeptides and/or the delivery of therapeutic nucleic acid molecules.  
 XX  
 SQ Sequence 60 BP; 13 A; 15 C; 17 G; 15 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7.16e-07 Length: 60  
 Score: 83.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 92.22% Indels: 0  
 DB: 18 Gaps: 0  
 US-09-771-009-1 (1-17) x AAT92990 (1-60)  
 QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGly 16  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 13 AATTTTGACCTTCTTAAGCTTGGCGGAGACGTCGAGTCCAAACCTGGG 60  
 RESULT 12  
 AAT92988  
 ID AAT92988 standard; DNA; 81 BP.  
 XX  
 AC AAT92988;  
 XX  
 DT 24-APR-1998 (first entry)  
 XX  
 DE Fragment from LITMUS 39 based plasmid pLit.GFP-2A23H-CP.  
 XX  
 KW Virus like particle; diagnostic reagent; antibiotic; therapy; VLP;  
 KW food supplement; vaccination; herbicide resistance; industrial enzyme;  
 KW tobacco mild green mosaic virus; TMGMV; PCR primer; LITMUS 39; ss.  
 XX  
 OS Synthetic.  
 OS Potato virus X.  
 XX  
 PN W09739134-A1.  
 XX  
 PD 23-OCT-1997.  
 XX  
 PF 17-APR-1997; 97WO-GB01065.  
 XX  
 PR 17-APR-1996; 96GB-0007899.  
 XX  
 PA (SCCR-) SCOTTISH CROP RES INST.  
 XX  
 PI Chapman SN, Wilson TMA;  
 XX  
 DR WPI; 1997-526468/48.  
 DR P-PSDB; AAW33910.

```

XX Production of virus like particles - using a nucleic acid sequence
PT capable of assembly with a protein having a first viral portion and
PT second non-viral portion
XX
XX Example 2; Fig 4; 33pp; English.
XX
CC This is a fragment from the LITMUS 39 based plasmid pLit.GFP-2A23H-CP.
CC These plasmids contain the green fluorescent protein (GFP)-2A- potato
CC virus X coat protein (CP) gene fusions. This is used as a source for
CC GFP-2A gene fusion with tobacco mild green mosaic virus (TMGMV) coat
CC protein (CP) gene and 3' UTR. This is used in a novel method for
CC producing a protein having a first (viral) portion and a second
CC (non-viral) portion. The method comprises, expressing the protein in a
CC cell, providing a nucleic acid sequence capable of assembly with the
CC protein into a virus-like particle (VLP), and permitting in vivo assembly
CC of the protein and nucleic acid into VLPs. The VLPs can be used for the
CC production of proteins such as diagnostic reagents, antibiotics,
CC therapeutic agents or food supplements. They can be used for e.g.
CC anti-botrytis agents, PGIPs, anti-insect Bacillus thuringiensis toxin and
CC herbicide resistance agents, industrial enzymes, pharmaceuticals,
CC therapeutic proteins, and nucleic acids, and as bioreactors. They can
CC also be used intact for presentation of peptide epitopes for vaccination
CC of animals, the production of therapeutic or industrial proteins and
CC polypeptides and/or the delivery of therapeutic nucleic acid molecules.
XX
SQ Sequence 81 BP; 18 A; 20 C; 22 G; 21 T; 0 other;

Alignment Scores:
Pred. No.: 1.02e-06 Length: 81
Score: 83.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.22% Indels: 0
DB: 18 Gaps: 0

US-09-771-009-1 (1-17) x AAT92988 (1-81)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGly 16
Db 34 AATTTTGACCTTCTTAAGCTTGGGGAGACGTCGAGTCCAACCCCTGGG 81

RESULT 13
AAT92991
ID AAT92991 standard; DNA; 183 BP.
XX
AC AAT92991;
XX
DT 24-APR-1998 (first entry)
XX
DE Fragment from LITMUS 39 based plasmid pLit.GFP-2A58K-CP.
XX
KW Virus like particle; diagnostic reagent; antibiotic; therapy; VLP;
KW food supplement; vaccination; herbicide resistance; industrial enzyme;
KW tobacco mild green mosaic virus; TMGMV; PCR primer; LITMUS 39; ss.
XX
OS Synthetic.
OS Potato virus X.
XX
XX WO9739134-A1.
XX
XX 23-OCT-1997.
XX
PF 17-APR-1997; 97WO-GB01065.
XX
PR 17-APR-1996; 96GB-0007899.
XX
PA (SCCR-) SCOTTISH CROP RES INST.
XX
PI Chapman SN, Wilson TMA;
XX

```

```

DR WPI; 1997-526468/48.
DR P-PSDB; AAW33912.
XX
PT Production of virus like particles - using a nucleic acid sequence
PT capable of assembly with a protein having a first viral portion and
PT second non-viral portion
XX
XX Example 2; Fig 4; 33pp; English.
XX
CC This is a fragment from the LITMUS 39 based plasmid pLit.GFP-2A58K-CP.
CC These plasmids contain the green fluorescent protein (GFP)-2A- potato
CC virus X coat protein (CP) gene fusions. This is used as a source for
CC GFP-2A gene fusion with tobacco mild green mosaic virus (TMGMV) coat
CC protein (CP) gene and 3' UTR. This is used in a novel method for
CC producing a protein having a first (viral) portion and a second
CC (non-viral) portion. The method comprises, expressing the protein in a
CC cell, providing a nucleic acid sequence capable of assembly with the
CC protein into a virus-like particle (VLP), and permitting in vivo assembly
CC of the protein and nucleic acid into VLPs. The VLPs can be used for the
CC production of proteins such as diagnostic reagents, antibiotics,
CC therapeutic agents or food supplements. They can be used for e.g.
CC anti-botrytis agents, PGIPs, anti-insect Bacillus thuringiensis toxin and
CC herbicide resistance agents, industrial enzymes, pharmaceuticals,
CC therapeutic proteins, and nucleic acids, and as bioreactors. They can
CC also be used intact for presentation of peptide epitopes for vaccination
CC of animals, the production of therapeutic or industrial proteins and
CC polypeptides and/or the delivery of therapeutic nucleic acid molecules.
XX
SQ Sequence 183 BP; 51 A; 48 C; 47 G; 37 T; 0 other;

Alignment Scores:
Pred. No.: 2.71e-06 Length: 183
Score: 83.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.22% Indels: 0
DB: 18 Gaps: 0

US-09-771-009-1 (1-17) x AAT92991 (1-183)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGly 16
Db 136 AATTTTGACCTTCTCAAGTTGGGGAGACGTCGAGTCCAACCCCTGGG 183

RESULT 14
AAT85178
ID AAT85178 standard; DNA; 7277 BP.
XX
AC AAT85178;
XX
DT 11-FEB-1998 (first entry)
XX
DE Equine rhinovirus 1 (ERHV1) polyprotein encoding nucleotide sequence.
XX
KW Equine rhinovirus 1; ERHV1; foot-and-mouth disease virus; vaccine; horse;
KW diagnosis; antigen; polyprotein; enzyme-linked immunosorbent assay;
KW recombinant protein; ss.
XX
OS Equine rhinovirus 1.
XX
XX Key Location/Qualifiers
XX CDS 435..7178
XX /*tag= a
XX FT /codon_start= putative
XX FT misc_feature 410..23
XX FT /*tag= b
XX FT /note= "polypyrimidine tract"
XX FT 5'UTR 1..434
XX FT /*tag= c
XX FT 3'UTR 7179..7277
XX FT /*tag= d

```



XX W09722701-A1.  
PN 26-JUN-1997.  
XX 18-DEC-1996; 96WO-AU00815.  
XX 18-DEC-1995; 95AU-0007201.  
XX (UYME ) UNIV MELBOURNE.  
XX Crabb BS, Feng L, Studdert MJ;  
PI WPI; 1997-341692/31.  
DR P-PSDB; AAW27126.  
XX Genomic sequence of equine rhinovirus 1 - and derived proteins or  
PT virus-like particles, useful in vaccines and as diagnostic agents  
XX Claim 1; Pages 27-29; 60pp; English.  
XX The present sequence represents the nucleic acid sequence encoding the  
CC polypeptide (AAW27126) of equine rhinovirus 1 (ERHV1). The taxonomic  
CC status of ERHV1 is unclear, as physicochemical studies have shown that  
CC the nucleic acid density and base composition of ERHV1 differs from other  
CC rhinoviruses. To this end, the nucleotide sequence encoding the  
CC polypeptide of ERHV1 was deduced. Analysis of this sequence suggests  
CC that ERHV1 is more closely related to foot-and-mouth disease virus  
CC than individual ERHV1 proteins can be used to make vaccines to protect horses  
CC (and possibly other animals) against ERHV1. Oligonucleotide primers and  
CC probes can be used for diagnosis of ERHV1 or related viruses, while  
CC antigens of the ERHV1 polypeptide can be used to detect ERHV1-specific  
CC antibodies in the blood, particularly in enzyme-linked immunosorbent  
CC assay. They can differentiate between infected animals and those  
CC vaccinated with ERHV1 vaccines (the infected animals will have antibodies  
CC reacting with non-capsid proteins but vaccinated animals will not).  
CC Fragments of the present sequence represent individual genes of the virus  
CC and can be expressed in host systems to produce recombinant proteins.  
CC Virus like particles containing the individual ERHV1 proteins, can also  
CC be used as vectors for delivering therapeutic or other useful agents,  
CC including vaccinating epitopes from other pathogens or reproductive  
CC hormones.  
XX SQ Sequence 7277 BP; 1805 A; 1662 C; 1750 G; 2060 T; 0 other;

Alignment Scores:  
Pred. No.: 0.000505 Length: 7277  
Score: 81.00 Matches: 15  
Percent Similarity: 94.12% Conservative: 1  
Best Local Similarity: 88.24% Mismatches: 1  
Query Match: 90.00% Indels: 0  
DB: 18 Gaps: 0

US-09-771-009-1 (1-17) x AAT85178 (1-7277)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
||||: |||||||||||||||||||||||||||||||||||||||||

Db 3408 AATTACTCTCTCCTCAATGGCTGGAGATGTTGAGAGCAACCTGGCCCC 3458

RESULT 15  
AAN20016  
ID AAN20016 standard; cDNA; 2802 BP.  
XX  
AC AAN20016;  
XX  
DT 20-AUG-1992 (first entry)  
XX  
DE Sequence encoding p20, VP4, VP2, VP3, VP1 and p52 in recombinant  
XX plasmid pFA A61/t 76.  
XX  
KW Vaccine; antibody; capsid protein; immunogen; antigen;  
XX foot and mouth disease; ss.

OS Foot and mouth disease virus.  
XX Key Location/Qualifiers  
FH CDS 2..226  
FT /\*tag= a  
FT /product= p20  
FT 227..527  
FT /\*tag= b  
FT /product= VP4  
FT /note= "claim 18"  
FT 528..1171  
FT /\*tag= c  
FT /product= VP2  
FT /note= "claim 16"  
FT 1172..1834  
FT /\*tag= d  
FT /product= VP3  
FT /note= "claim 17"  
FT 1835..2470  
FT /\*tag= e  
FT /product= VP1  
FT /note= "claim 15"  
FT 2471..2802  
FT /\*tag= f  
FT /product= p52  
XX EP48455-A.  
XX 31-MAR-1982.  
XX 18-SEP-1980; 80GB-0030208.  
PR 22-OCT-1980; 80GB-0034130.  
PR 27-NOV-1980; 80GB-0038147.  
PR 08-APR-1981; 81GB-0011064.  
PR 18-AUG-1981; 81GB-0025150.  
XX (NATR ) NATIONAL RES DEV CORP.  
PA (WELL ) WELLCOME FOUNDATION LTD.  
XX Boothroyd JC, Cross GAM, Highfield PE, Winther MD, Rowlands DJ;  
PI Brown F, Harris TJR, Lowe PA;  
XX WPI; 1982-26702E/14.  
DR P-PSDB; AAP20016.  
XX DNA corresp. to (part of) foot and mouth disease virus RNA - useful  
PT in prepn. of vaccines for producing antibodies against the virus  
XX Disclosure; Fig 13; 57pp; English.  
CC The inventors claim a DNA molecule comprising a nucleotide sequence  
CC corresp. to all or a portion of foot-and-mouth disease virus RNA  
CC (FMDV). The DNA molecule is esp. for a precursor of FMDV capsid  
CC protein. It esp. codes for FMDV protein p88 and VP1-VP4. It may code  
CC for VP4, VP2, VP3 and VP1 contiguously. The inventors also claim a  
CC vaccine for stimulating prodn. of antibodies against FMDV in a  
CC mammal which comprises at least one of the above recombinant  
CC proteins produced by a host cell transformed with the DNA.  
XX SQ Sequence 2802 BP; 714 A; 810 C; 689 G; 588 T; 0 other;

Alignment Scores:  
Pred. No.: 0.000245 Length: 2802  
Score: 80.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 88.89% Indels: 0  
DB: 3 Gaps: 0

US-09-771-009-1 (1-17) x AAN20016 (1-2802)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
|||||||||||||||||||||||||||||||||||||||||||

Db 2471 AACCTTGACCTACTTAAGTTGGCGGTGACGTTGAGTCCAAACCTTGGGCC 2521

Search completed: January 24, 2003, 18:45:55  
Job time : 305 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 24, 2003, 18:39:14 ; Search time 3166 Seconds  
(without alignments)  
156.269 Million cell updates/sec

Title: US-09-771-b09-1  
Perfect score: 90  
Sequence: 1 NFDLLKLAGDVESNFGP 17

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DRV=xlp  
-Q/cgn2\_1/USPTO.spool/US09771009/runat\_22012003\_140100\_18450/app\_query.fasta\_1.199  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09771009.ecgn\_1.1.3637@runat\_22012003\_140100\_18450 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	90	100.0	327	14	FMDVAICP1 X8855 Foot and mo
2	90	100.0	327	14	FMDVAICP2 X8856 Foot and mo
3	90	100.0	327	14	FMDVAICP3 X8857 Foot and mo
4	90	100.0	330	14	FMDVOCIP2 X8862 Foot and mo
5	90	100.0	330	14	FMDVA22C2 X8860 Foot and mo
6	90	100.0	330	14	FMDVA22CP X8859 Foot and mo
7	90	100.0	330	14	FMDVAICP4 X8858 Foot and mo
8	90	100.0	330	14	FMDVOCIP1 X8861 Foot and mo
9	90	100.0	330	14	FMDVOCIP3 X8863 Foot and mo
10	90	100.0	493	14	AF024509 AF024509 Foot-and-
11	90	100.0	750	14	APHOVLIN M16078 Foot-and-
12	90	100.0	873	6	AX364557 AX364557 Sequence
13	90	100.0	939	6	AX403959 AX403959 Sequence
14	90	100.0	939	6	AX460903 AX460903 Sequence
15	90	100.0	961	6	A00276 A00276 Foot and mo
16	90	100.0	961	6	E00030 E00030 DNA coding
17	90	100.0	1105	6	AX364441 AX364441 Sequence
18	90	100.0	1163	14	PIFMDV V01131 Foot-and mo
19	90	100.0	3119	14	PIFMDC X00130 Foot-and-mo
20	90	100.0	3575	14	APHPVP M95781 Foot and mo
21	90	100.0	3579	14	FDI251476 AJ251476 Foot-and-
22	90	100.0	5713	6	AX146811 AX146811 Sequence
23	90	100.0	7733	14	AF026168 AF026168 Foot-and-
24	90	100.0	7739	14	AF154271 AF154271 Foot-and-
25	90	100.0	7774	14	FDI251473 FDI251473 Foot-and-
26	90	100.0	7804	14	PIFMDY2 X00871 Foot and mo
27	90	100.0	7820	14	FMDVALF X74812 Foot and Mo
28	90	100.0	8115	14	AF274010 AF274010 Foot-and-
29	90	100.0	8115	14	FAN13358 AJ133358 Foot-and-
30	90	100.0	8115	14	FAN13359 AJ133359 Foot-and-
31	90	100.0	8115	14	FDI133357 AJ133357 Foot-and-
32	90	100.0	8134	14	AF308157 AF308157 Foot-and-
33	90	100.0	8147	14	AF511039 AF511039 Foot-and-
34	90	100.0	8161	14	FMV7347 AJ007347 Foot-and-
35	90	100.0	8161	14	FMV7572 AJ007572 Foot-and-
36	90	100.0	8168	14	FDI320488 AJ320488 Foot-and-
37	90	100.0	8173	14	AF506822 AF506822 Foot-and-
38	90	100.0	8821	14	AF361253 AF361253 Equine rh
39	87	96.7	6996	14	AF189157 AF189157 Foot-and-
40	85	94.4	401	14	AF167307 AF167307 Foot-and-
41	85	94.4	7813	14	AF377945 AF377945 Foot-and-
42	85	94.4	7822	14	AB079061 AB079061 Foot-and-
43	83	92.2	60	6	A66682 A66682 Sequence 12
44	83	92.2	60	6	A66684 A66684 Sequence 14
45	83	92.2	60	6	AR151648 AR151648 Sequence

ALIGNMENTS

```

FMDVAICP1
LOCUS       327 bp      mRNA      linear      VRL 31-AUG-1995
DEFINITION  Foot and mouth disease virus Asia I mRNA for capsid protein
            (partial) and protease (Nainital).
ACCESSION   X88855
VERSION     X88855.1 GI:971405
KEYWORDS    2a protease; 2B protease; immunogenic protein; polyprotein; VP1
            gene; VP1 protein.
SOURCE      Foot-and-mouth disease virus.
            Picornaviridae: Aphthovirus.
ORGANISM    Foot-and-mouth disease virus.
            Viruses: ssRNA positive-strand viruses, no DNA stage;
            Picornaviridae: Aphthovirus.
REFERENCE   1 (bases 1 to 327)
AUTHORS     Tulasiram,P., Tyagi,M. and Suryanarayana,V.
TITLE       Antigenic variation in foot and mouth disease virus type Asia I
            isolates of India
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 327)
AUTHORS     Suryanarayana,V.V.S.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary
            Research Institute, Scientist (SG), Hebbal, Bangalore 560 024,
            INDIA

FEATURES             Location/Qualifiers
     source          1..327
                     /organism="Foot-and-mouth disease virus"
                     /isolate="Asia I Nainital, U.P. (vaccine strain)"
                     /db_xref="taxon:12110"
     CDS             /map="3337-3669 of 1D with 2A of FMDV"
                     /cell_line="Baby Hamster Kidney (BHK)21 clone 13"
                     <1..>327
                     /codon_start=1
                     /product="immunogenic polyprotein with 2A protease"
                     /protein_id="CAA61325.1"
                     /db_xref="GI:971406"
                     /db_xref="SPTREMBL:Q67450"
                     /translation="APHRVLATVYNGKTYGTPRRGDLAVLAQRVSNRLPTSFNYG
                     AVKADITELLIRMKRAETCYPRLLALDTHDRRKQIIAIAPEKQVLFNFDLLKLAGDV
                     ESNPGPF"
     mat_peptide     <1..273
                     /gene="VP1"
                     /product="VP1 protein"
                     274..321
                     /product="2A protease"
                     322..>327
                     /product="2B protease"
     gene            1..273
                     /gene="VP1"

BASE COUNT      76 a      91 g      62 t
ORIGIN
Alignment Scores:
Pred. No.:      2.44e-08      Length:      327
Score:          90.00      Matches:      17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             14      Gaps:      0

US-09-771-009-1 (1-17) x FMDVAICP1 (1-327)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
|||||
Db 274 AACTTTGACCTGCTGAAGTTGGCGGAGACGTGGAGTCCAACCTGGGCC 324

RESULT 2
FMDVAICP2
LOCUS       327 bp      mRNA      linear      VRL 31-AUG-1995
DEFINITION  Foot and mouth disease virus Asia I mRNA for capsid protein
            (partial) and protease (Mandya).
ACCESSION   X88856
VERSION     X88856.1 GI:971407
KEYWORDS    2a protease; 2B protease; immunogenic protein; polyprotein; VP1
            gene; VP1 protein.
SOURCE      Foot-and-mouth disease virus.
            Picornaviridae: Aphthovirus.
ORGANISM    Foot-and-mouth disease virus.
            Viruses: ssRNA positive-strand viruses, no DNA stage;
            Picornaviridae: Aphthovirus.
REFERENCE   1 (bases 1 to 327)
AUTHORS     Tulasiram,P., Tyagi,M. and Suryanarayana,V.
TITLE       Antigenic variation in foot and mouth disease virus type Asia I
            isolates of India
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 327)
AUTHORS     Suryanarayana,V.V.S.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary
            Research Institute, Scientist (SG), Hebbal, Bangalore 560 024,
            INDIA

FEATURES             Location/Qualifiers
     source          1..327
                     /organism="Foot-and-mouth disease virus"
                     /isolate="Asia I Nainital, U.P. (vaccine strain)"
                     /db_xref="taxon:12110"
     CDS             /map="3337-3669 of 1D with 2A of FMDV"
                     /cell_line="Baby Hamster Kidney (BHK)21 clone 13"
                     <1..>327
                     /codon_start=1
                     /product="immunogenic polyprotein with 2A protease"
                     /protein_id="CAA61325.1"
                     /db_xref="GI:971406"
                     /db_xref="SPTREMBL:Q67450"
                     /translation="APHRVLATVYNGKTYGTPRRGDLAVLAQRVSNRLPTSFNYG
                     AVKADITELLIRMKRAETCYPRLLALDTHDRRKQIIAIAPEKQVLFNFDLLKLAGDV
                     ESNPGPF"
     mat_peptide     <1..273
                     /gene="VP1"
                     /product="VP1 protein"
                     274..321
                     /product="2A protease"
                     322..>327
                     /product="2B protease"
     gene            1..273
                     /gene="VP1"

BASE COUNT      76 a      91 g      62 t
ORIGIN
Alignment Scores:
Pred. No.:      2.44e-08      Length:      327
Score:          90.00      Matches:      17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             14      Gaps:      0

US-09-771-009-1 (1-17) x FMDVAICP1 (1-327)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
|||||
Db 274 AACTTTGACCTGCTGAAGTTGGCGGAGACGTGGAGTCCAACCTGGGCC 324

RESULT 3
FMDVAICP3
LOCUS       327 bp      mRNA      linear      VRL 31-AUG-1995
DEFINITION  Foot and mouth disease virus Asia I mRNA for capsid protein
            (partial) and protease (Nilgiri).
ACCESSION   X88857
VERSION     X88857.1 GI:971409
KEYWORDS    2a protease; 2B protease; immunogenic protein; polyprotein; VP1
            gene; VP1 protein.
SOURCE      Foot-and-mouth disease virus.
            Picornaviridae: Aphthovirus.
ORGANISM    Foot-and-mouth disease virus.
            Viruses: ssRNA positive-strand viruses, no DNA stage;
            Picornaviridae: Aphthovirus.
REFERENCE   1 (bases 1 to 327)
AUTHORS     Tulasiram,P., Tyagi,M. and Suryanarayana,V.
TITLE       Antigenic variation in foot and mouth disease virus type Asia I
            isolates of India
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 327)
AUTHORS     Suryanarayana,V.V.S.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary
            Research Institute, Scientist (SG), Hebbal, Bangalore 560 024,
            INDIA

FEATURES             Location/Qualifiers
     source          1..327
                     /organism="Foot-and-mouth disease virus"
                     /isolate="Asia I Mandya, Karnataka"
                     /db_xref="taxon:12110"
     CDS             /map="3337-3669 of 1D with 2A of FMDV"
                     <1..>327
                     /codon_start=1
                     /product="immunogenic polyprotein with 2A protease"
                     /protein_id="CAA61326.1"
                     /db_xref="GI:971408"
                     /db_xref="SPTREMBL:Q67451"
                     /translation="APHRVLATVYNGKTYGTPRRGDMALAQRLSRQLPTSFNYG
                     AVKAENITELLIRMKRAETCYPRLLALDTHDRRKQIIAIAPEKQVLFNFDLLKLAGDV
                     ESNPGPF"
     mat_peptide     <1..273
                     /gene="VP1"
                     /product="VP1 protein"
                     274..321
                     /product="2A protease"
                     322..>327
                     /product="2B protease"
     gene            1..273
                     /gene="VP1"

BASE COUNT      77 a      96 c      93 g      61 t
ORIGIN
Alignment Scores:
Pred. No.:      2.44e-08      Length:      327
Score:          90.00      Matches:      17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             14      Gaps:      0

US-09-771-009-1 (1-17) x FMDVAICP2 (1-327)

QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
|||||
Db 274 AACTTTGACCTGCTGAAGTTGGCGGAGACGTGGAGTCCAACCTGGGCC 324

RESULT 3
FMDVAICP3
LOCUS       327 bp      mRNA      linear      VRL 31-AUG-1995
DEFINITION  Foot and mouth disease virus Asia I mRNA for capsid protein
            (partial) and protease (Nilgiri).
ACCESSION   X88857
VERSION     X88857.1 GI:971409
KEYWORDS    2a protease; 2B protease; immunogenic protein; polyprotein; VP1
            gene; VP1 protein.
SOURCE      Foot-and-mouth disease virus.
            Picornaviridae: Aphthovirus.
ORGANISM    Foot-and-mouth disease virus.
            Viruses: ssRNA positive-strand viruses, no DNA stage;
            Picornaviridae: Aphthovirus.
REFERENCE   1 (bases 1 to 327)
AUTHORS     Tulasiram,P., Tyagi,M. and Suryanarayana,V.
TITLE       Antigenic variation in foot and mouth disease virus type Asia I
            isolates of India
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 327)
AUTHORS     Suryanarayana,V.V.S.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary
            Research Institute, Scientist (SG), Hebbal, Bangalore 560 024,
            INDIA

```

Isolates of India  
 Unpublished  
 REFERENCE 2 (bases 1 to 327)  
 AUTHORS Suryanarayana, V.V.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary Research Institute, Scientist (SG), Hebbal, Bangalore 560 024, INDIA

**FEATURES**  
 Location/Qualifiers  
 source 1..327  
 /organism="Foot-and-mouth disease virus"  
 /isolate="Asia I Nilgiri, Tamil Nadu"  
 /db\_xref="taxon:12110"  
 /map="3337-3669 of 1D with 2A of FMDV"  
 <1..>327  
 /codon\_start=1  
 /product="immunogenic polypeptide with 2A protease"  
 /protein\_id="CAA61327.1"  
 /db\_xref="GI:971410"  
 /db\_xref="SPTREMBL:Q67452"  
 /translation="APHRVLATVYNGKTYGEQPSRRGDMAALQRLSROLPTSFNYG AVKAENITELLIRMKRAETPCPRLALDTAQRKOEIIAPKKQVLNFDLLKLGVDV ESNPGPF"  
 mat\_peptide <1..273  
 /gene="vp1"  
 /product="vp1 protein"  
 274..321  
 mat\_peptide /product="2A protease"  
 322..>327  
 mat\_peptide /product="2B protease"  
 1..273  
 gene /gene="vp1"  
 78 a 96 c 93 g 60 t

**BASE COUNT** 78 a 96 c 93 g 60 t  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 2.44e-08 Length: 327  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

US-09-771-009-1 (1-17) x FMDVAICP3 (1-327)  
 QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 274 AACTTTCACCTGCTGAAGTTGGCGGAGACGTGGAGTCCACCTGGGCC 324

**RESULT 4**  
 FMDV0CP02 FMDV0CP02 330 bp mRNA linear VRL 07-JAN-1997  
 LOCUS Foot and mouth disease virus 0 mRNA for capsid protein (partial)  
 DEFINITION and protease (OEK).  
 ACCESSION X88862  
 VERSION X88862.1 GI:971399  
 KEYWORDS 2a protease; 2B protease; immunogenic protein; polyprotein; VP1  
 gene; vp1 protein.  
 SOURCE Foot-and-mouth disease virus.  
 ORGANISM Foot-and-mouth disease virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Picornaviridae; Aphthovirus.  
 REFERENCE 1 (bases 1 to 330)  
 AUTHORS Suryanarayana, V.V.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary Research Institute, Scientist (SG), Hebbal, Bangalore 560 024, INDIA  
 REFERENCE 2 (bases 1 to 330)  
 AUTHORS Singh, M., Mohan, B.M. and Suryanarayana, V.V.  
 TITLE Serological and molecular analysis of serotype O foot-and-mouth disease virus isolated from disease outbreaks in India during 1987-91

Virus Res. 43 (1), 45-55 (1996)  
 MEDLINE 96419934  
 PUBMED 8822633  
**FEATURES**  
 Location/Qualifiers  
 source 1..330  
 /organism="Foot-and-mouth disease virus"  
 /isolate="type O OEK"  
 /db\_xref="taxon:12110"  
 /map="3337-3669 of 1D with 2A of FMDV"  
 <1..>330  
 /codon\_start=1  
 /product="immunogenic polypeptide with 2A protease"  
 /protein\_id="CAA61332.1"  
 /db\_xref="GI:971400"  
 /db\_xref="SPTREMBL:Q67432"  
 /translation="APHRVLATVYNGKTYGEQPSRRGDMAALQRLSROLPTSFNYG YGAKATGVTELPYRMPRAETPCPRLALDTHPSEARNHEDCATLEQLNFDLLKLGADV VESNPGPF"  
 mat\_peptide <1..276  
 /gene="vp1"  
 /product="vp1 protein"  
 277..324  
 mat\_peptide /product="2A protease"  
 325..>330  
 mat\_peptide /product="2B protease"  
 1..276  
 gene /gene="vp1"  
 75 a 100 c 90 g 65 t

**BASE COUNT** 75 a 100 c 90 g 65 t  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 2.46e-08 Length: 330  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

US-09-771-009-1 (1-17) x FMDV0CP02 (1-330)  
 QY 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 277 AACTTCGACCTGCTCAAGTTGGCGGAGACGTGGAGTCCACCTGGGCC 327

**RESULT 5**  
 FMDVA22C2 FMDVA22C2 330 bp mRNA linear VRL 31-AUG-1995  
 LOCUS Foot and mouth disease virus A22 mRNA for capsid protein (partial)  
 DEFINITION and protease (Tamil Nadu).  
 ACCESSION X88860  
 VERSION X88860.1 GI:971401  
 KEYWORDS 2a protease; 2B protease; immunogenic protein; polyprotein; VP1  
 gene; vp1 protein.  
 SOURCE Foot-and-mouth disease virus.  
 ORGANISM Foot-and-mouth disease virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Picornaviridae; Aphthovirus.  
 REFERENCE 1 (bases 1 to 330)  
 AUTHORS Tulasiram, P., Tyagi, M. and Suryanarayana, V.  
 TITLE Antigenic variation in foot and mouth disease virus type Asia I isolates of India  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 330)  
 AUTHORS Suryanarayana, V.V.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary Research Institute, Scientist (SG), Hebbal, Bangalore 560 024, INDIA  
**FEATURES**  
 Location/Qualifiers  
 source 1..330  
 /organism="Foot-and-mouth disease virus"  
 /isolate="A22 Tamil Nadu, India"  
 /db\_xref="taxon:12110"



```

mat_peptide 325..>330
gene 1..276
BASE COUNT 74 a 105 c 88 g 63 t
ORIGIN
Alignment Scores:
Pred. No.: 2.46e-08 Length: 330
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-771-009-1 (1-17) x FMDVAICP4 (1-330)

Qy 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
Db 277 AACTTTGACCTACTTAAGTTGGCGGAGACGTTGAGTCCAAACCCCTGGGCC 327

RESULT 8
FMDVOC01
LOCUS FMDVOC01 330 bp mRNA linear VRL 07-JAN-1997
DEFINITION Foot and mouth disease virus 0 mRNA for capsid protein (partial)
and protease (Karnataka).
ACCESSION X88861
VERSION X88861.1 GI:971413
KEYWORDS 2a protease; 2B protease; immunogenic protein; polyprotein; VP1
gene; VP1 protein.
SOURCE Foot-and-mouth disease virus.
ORGANISM Foot-and-mouth disease virus.
REFERENCE 1 (bases 1 to 330)
AUTHORS Suryanarayana,V.V.S.
TITLES Direct Submission
JOURNAL Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary
Research Institute, Scientist (SG), Hebbal, Bangalore 560 024,
INDIA

REFERENCE 2 (bases 1 to 330)
AUTHORS Singh,M., Mohan,B.M. and Suryanarayana,V.V.
TITLES Serological and molecular analysis of serotype O foot-and-mouth
disease virus isolated from disease outbreaks in India during
1987-91
JOURNAL Virus Res. 43 (1), 45-55 (1996)
MEDLINE 96419934
PUBMED 8822633
FEATURES
source
Location/Qualifiers
1..330
/organism="Foot-and-mouth disease virus"
/isolate="type O Karnataka, India"
/db_xref="taxon:12110"
/map="3337-3669 of 1D with 2A of FMDV"
<1..>330
/codon_start=1
/product="immunogenic polyprotein with 2A protease"
/protein_id="CAA61331.1"
/db_xref="GI:971414"
/db_xref="SPTREMBL:O67454"
/translation="APHRLVATYNGNCKYGDGVTNIRGDDQVLAQKAARALPTSFN
YGAIKATRVTELLYRMKRAETPCRPILLADIPNEARHKEDCAPVKQLNFDLLKLAGD
VESNPGPF"
<1..276
/gene="VP1"
/mat_peptide 277..324
/mat_peptide 325..>330
/mat_peptide 1..276
gene
BASE COUNT 80 a 93 c 94 g 63 t
/gene="VP1"

mat_peptide 77 a 94 c 95 g 64 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 2.46e-08 Length: 330
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-771-009-1 (1-17) x FMDVOC01 (1-330)

Qy 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
Db 277 AACTTTGACCTGCTCAAGTTGGCGGAGACGTTGAGTCCAAACCCCTGGGCC 327

RESULT 9
FMDVOC03
LOCUS FMDVOC03 330 bp mRNA linear VRL 07-JAN-1997
DEFINITION Foot and mouth disease virus 0 mRNA for capsid protein (partial)
and protease (vaccine strain).
ACCESSION X88863
VERSION X88863.1 GI:971415
KEYWORDS 2a protease; 2B protease; immunogenic protein; polyprotein; VP1
gene; VP1 protein.
SOURCE Foot-and-mouth disease virus.
ORGANISM Foot-and-mouth disease virus.
REFERENCE 1 (bases 1 to 330)
AUTHORS Suryanarayana,V.V.S.
TITLES Direct Submission
JOURNAL Submitted (20-JUN-1995) V.V.S. Suryanarayana, Indian Veterinary
Research Institute, Scientist (SG), Hebbal, Bangalore 560 024,
INDIA

REFERENCE 2 (bases 1 to 330)
AUTHORS Singh,M., Mohan,B.M. and Suryanarayana,V.V.
TITLES Serological and molecular analysis of serotype O foot-and-mouth
disease virus isolated from disease outbreaks in India during
1987-91
JOURNAL Virus Res. 43 (1), 45-55 (1996)
MEDLINE 96419934
PUBMED 8822633
FEATURES
source
Location/Qualifiers
1..330
/organism="Foot-and-mouth disease virus"
/isolate="type O vaccine (O/R2/75)"
/db_xref="taxon:12110"
/map="3337-3669 of 1D with 2A of FMDV"
<1..>330
/codon_start=1
/product="immunogenic polyprotein with 2A protease"
/protein_id="CAA61333.1"
/db_xref="GI:971416"
/db_xref="SPTREMBL:O67455"
/translation="APHRLVATYNGNCKYGDGAVTNIRGDLQVLAQKAARALPTSEN
YGAIKATRVTELLYRMKRAETPCRPILLAIHPNEARNEDCAPVKQLNFDLLKLAGD
VESNPGPF"
<1..276
/gene="VP1"
/mat_peptide 277..324
/mat_peptide 325..>330
/mat_peptide 1..276
gene
BASE COUNT 77 a 94 c 95 g 64 t
/gene="VP1"

mat_peptide 77 a 94 c 95 g 64 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 2.46e-08 Length: 330
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-771-009-1 (1-17) x FMDVOC01 (1-330)

Qy 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17
Db 277 AACTTTGACCTGCTCAAGTTGGCGGAGACGTTGAGTCCAAACCCCTGGGCC 327

```

Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

US-09-771-009-1 (1-17) x FMDVCP03 (1-330)

Qy 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 |||||  
 Db 277 AACTTTGACCTGCTCAAGTTGGCGGAGACGTGGAGTCCAAACCTGGGCC 327

RESULT 10

AF024509  
 LOCUS AF024509 493 bp RNA linear VRL 26-MAR-1998  
 DEFINITION Foot-and-mouth disease virus Asia-1 polyprotein gene, partial cds.  
 ACCESSION AF024509  
 VERSION AF024509.1 GI:2988470

KEYWORDS  
 ORGANISM

Foot-and-mouth disease virus.  
 Foot-and-mouth disease virus  
 Viruses: ssRNA positive-strand viruses, no DNA stage;  
 Picornaviridae; Aphthovirus.

REFERENCE 1 (bases 1 to 493)

AUTHORS Bastos,A.D.  
 TITLE Detection and characterization of foot-and-mouth disease virus in

JOURNAL sub-Saharan Africa

MEDLINE Onderstepoort J. Vet. Res. 65 (1), 37-47 (1998)

PUBMED 98293234

REFERENCE 2 (bases 1 to 493)

AUTHORS Bastos,A.D.S.

TITLE Direct Submission

JOURNAL Submitted (10-SEP-1997) Research and Diagnostics, Onderstepoort

Institute for Exotic Diseases, Ou Soutpan Rd, Onderstepoort,

Pretoria, Gauteng 0110, South Africa

FEATURES

source

1..493 /organism="Foot-and-mouth disease virus"

/strain="Asia-1"

/isolate="PAK 1/54"

/db\_xref="taxon:12110"

<1..>493

/codon\_start=3

/product="polyprotein"

/protein\_id="AAC40617.1"

/db\_xref="GI:2988471"

/translation="LVHTGPTVWVPGAPKALDQNTNPTAYHKQPTIRALPYTAPH

RVLATVYNGTKTTEGPTMRGDCAVLASKYVKQLPTSFNGAYKAEINTEMLIRKRA

ETVCPRLALDTQDRKQEIITAEKQLLFOLLKLAGDVESNPGPFPTSDVRSNXT

KLV"

BASE COUNT 120 a 148 c 127 g 97 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 3.87e-08 Length: 493

Score: 90.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0

US-09-771-009-1 (1-17) x AF024509 (1-493)

Qy 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17

|||||

Db 399 AACTTTGACCTACTCAAGTTGGCGGAGACGTGGAGTCCAAACCTGGGCC 449

RESULT 11

LOCUS APOHVP1N 750 bp ss-RNA linear VRL 28-APR-1993

DEFINITION Foot and mouth disease virus capsid protein VP1 RNA, complete cds.

ACCESSION M16078

MI6078.1 GI:210468  
 KEYWORDS vp1 capsid protein.  
 SOURCE Foot and mouth disease virus, cDNA to viral RNA.  
 ORGANISM Foot-and-mouth disease virus

Viruses: ssRNA positive-strand viruses, no DNA stage;  
 Picornaviridae; Aphthovirus.

REFERENCE 1 (bases 1 to 750)

AUTHORS Beck,E. and Strohmaier,K.

TITLE Subtyping of European foot-and-mouth disease virus strains by

Journal nucleotide sequence determination

Journal J. Virol. 61 (5), 1621-1629 (1987)

MEDLINE 87198892

PUBMED 3033288

FEATURES

source

1..750

/organism="Foot-and-mouth disease virus"

/db\_xref="taxon:12110"

1..639

/note="VP1 capsid protein"

BASE COUNT 198 a 223 c 180 g 149 t

ORIGIN

Alignment Scores:

Pred. No.: 6.21e-08 Length: 750

Score: 90.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0

US-09-771-009-1 (1-17) x APOHVP1N (1-750)

Qy 1 AsnPheAspLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17

|||||

Db 640 AATTTGACCTTCTCAAGTTGGCGGAGACGTGAGTCCAAACCTGGGCC 690

RESULT 12

AX364557

LOCUS AX364557 873 bp DNA linear PAT 15-FEB-2001

DEFINITION Sequence 564 from Patent WO0208410.

ACCESSION AX364557

VERSION AX364557.1 GI:18696517

KEYWORDS

SOURCE

synthetic construct.

synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS

Reddy,S.A., Larrinua,M.I., Ruegger,M., Weglarz,T., Blakeslee,B.,

Oriedo,V.B., Savickas,J.P., McCreary,A.D., Miller,A.B., Poque,P.G.,

Della-Cioppa,R.G., Wolfe,M.G., Zheng,W., Gachotte,D., Grosley,R.

and Pell,R.

TITLE Nucleic acids compositions conferring dwarfing phenotype

JOURNAL

Patent: WO 0208410-A 564 31-JAN-2002;

The Dow Chemical Company (US) ; Reddy, S. Avutu (US) ; Larrinua, M.

Ignacio (US) ; Ruegger, Max (US) ; Weglarz, Ted (US) ; Blakeslee,

Beth (US)

FEATURES

source

1..873

/organism="synthetic construct"

/db\_xref="taxon:32630"

BASE COUNT 201 a 203 c 239 g 230 t

ORIGIN

Alignment Scores:

Pred. No.: 7.37e-08 Length: 873

Score: 90.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-771-009-1 (1-17) x AX364557 (1-873)



QY 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 |||||||  
 Db 686 AATTTTGATTGCTGAAGTTGGGGGTGACGTGGAATCTCAACCCCTGGTCT 736

RESULT 13  
 AX403959  
 LOCUS AX403959 939 bp DNA linear PAT 14-JUN-2002  
 DEFINITION Sequence 7 from Patent EP1195438.  
 ACCESSION AX403959  
 VERSION AX403959.1 GI:21437305

KEYWORDS synthetic construct.  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.

REFERENCE 1  
 AUTHORS Regts,D.G., Wilschut,J.C., Holtrop,M. and Daemen,C.A.  
 TITLE Genetic immunisation against cervical carcinoma  
 JOURNAL Patent: EP 1195438-A 7 10-APR-2002;  
 Rijksuniversiteit te Groningen (NL)

FEATURES  
 Location/Qualifiers  
 1..939  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="nucleotide sequence of construct enh E6, 7"  
 BASE COUNT 284 a 197 c 217 g 241 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8e-08 Length: 939  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-771-009-1 (1-17) x AX403959 (1-939)

QY 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 |||||||  
 Db 114 AATTTTGACCTCTTAAGCTTGGGGGAGACGTGAGTCCAACCCCTGGGCC 164

RESULT 14  
 AX460903  
 LOCUS AX460903 939 bp DNA linear PAT 08-JUL-2002  
 DEFINITION Sequence 7 from Patent WO0229074.  
 ACCESSION AX460903  
 VERSION AX460903.1 GI:21726146  
 KEYWORDS synthetic construct.  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.

REFERENCE 1  
 AUTHORS Regts,D.G., Holtrop,M., Wilschut,J.C. and Daemen,C.A.  
 TITLE Genetic immunisation against cervical carcinoma  
 JOURNAL Patent: WO 0229074-A 7 11-APR-2002;  
 Rijksuniversiteit Groningen (NL)

FEATURES  
 Location/Qualifiers  
 1..939  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="nucleotide sequence of construct enh E6, 7"  
 BASE COUNT 284 a 197 c 217 g 241 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8e-08 Length: 939  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-771-009-1 (1-17) x AX460903 (1-939)

QY 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 |||||||  
 Db 114 AATTTTGACCTCTTAAGCTTGGGGGAGACGTGAGTCCAACCCCTGGGCC 164

RESULT 15  
 A00276  
 LOCUS A00276 961 bp DNA linear PAT 14-JUN-1995  
 DEFINITION Foot and mouth disease virus transgenic DNA.  
 ACCESSION A00276  
 VERSION A00276.1 GI:14462

KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE 1  
 AUTHORS Nofschneider,P.H., Schaller,H., Kupper,H.A. and Keller.  
 JOURNAL Patent: GB 2079288-A 2 20-JAN-1982;  
 Biogen N.V

FEATURES  
 Location/Qualifiers  
 1..961  
 /organism="Foot-and-mouth disease virus"  
 /db\_xref="taxon:12110"  
 1..961  
 /gene="vp1"  
 <1..>961  
 /gene="VP1"  
 /note="DNA insert FMDV-1034"  
 /codon\_start=2  
 /protein\_id="CAA00045.1"  
 /db\_xref="GI:1334776"

ORIGIN  
 BASE COUNT 251 a 278 c 240 g 192 t  
 Alignment Scores:  
 Pred. No.: 8.21e-08 Length: 961  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-771-009-1 (1-17) x A00276 (1-961)

QY 1 AsnPheAspLeuLeuLysLeuAlaGlyAspValGluSerAsnProGlyPro 17  
 |||||||  
 Db 701 AATTTTGACCTCTCAAGTTGGGGGAGACGTGAGTCCAACCCCTGGGCC 751

Search completed: January 24, 2003, 19:38:53  
 Job time : 3169 secs

**THIS PAGE BLANK (USPTO)**